PREDICTABILITY AND PREDICTION OF TROPICAL CYCLONES ON DAILY TO INTERANNUAL TIME SCALES

A Dissertation
Presented to
The Academic Faculty

by

James Ian Belanger

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Earth and Atmospheric Sciences

Georgia Institute of Technology
August 2012

PREDICTABILITY AND PREDICTION OF TROPICAL CYCLONES ON DAILY TO INTERANNUAL TIME SCALES

Approved by:

Dr. Judith A. Curry, Advisor
School of Earth and Atmospheric Sciences
Georgia Institute of Technology

Dr. Yi Deng
School of Earth and Atmospheric Sciences
Georgia Institute of Technology

Dr. Peter J. Webster
School of Earth and Atmospheric Sciences
Georgia Institute of Technology

Dr. Greg J. Holland
The Mesoscale and Microscale Meteorology Division
NCAR Earth System Laboratory

Dr. Robert X. Black
School of Earth and Atmospheric Sciences
Georgia Institute of Technology

Date Approved:

To the more than 138,000 men, women, and children who unnecessarily lost their lives as Cyclone Nargis moved inland from the North Indian Ocean. May we learn from their death the true value in extended-range tropical cyclone forecasts.

ACKNOWLEDGEMENTS

This dissertation reflects the unending love, support and sacrifice of several influential people in my life. First, I give thanks to my heavenly Father, from whom all blessings flow, for a Faith that has allowed me to reach this point today. From a scientific perspective, I would like to thank my advisor Dr. Judith A. Curry and co-advisor Dr. Peter J. Webster. They have provided a slew of opportunities and have given me sufficient independence to pursue a rich set of interesting research problems. In addition, their early recognition for my affinity to tropical cyclones has paved the way for a variety of projects related to assessing and improving weather predictions both through Georgia Tech and CFAN. They have also been important role models both in their commitment to scientific integrity and in their dedication to high ethical standards that will serve me well as a future research scientist. In addition, I thank my committee members, Dr. Robert X. Black and Dr. Yi Deng for their review of my dissertation and for their scientific advice during my time at Georgia Tech. I am also grateful for my external committee member, Dr. Greg J. Holland. Dr. Holland's passion for tropical cyclones has led to an exhaustive research publication record, which I have found extremely helpful in shaping my views in this field. In addition, he has provided several opportunities for scientific growth through collaborative research projects and through engaging visits to the National Center for Atmospheric Research. I am also grateful to have had the opportunity to spend the first year of my graduate career at SUNY Albany. The course work under Dr. Chris D. Thorncroft and Dr. Paul Roundy along with research opportunities under Dr. John Molinari provided essential background knowledge in tropical atmospheric dynamics. Finally, I must give special recognition to Mark T. Jelinek and Violeta E. Toma both on a personal and professional level for innumerable discussions and for their sage advice through the dissertation process.

On a personal note, I would like to thank my parents for their love and sacrifice that they have made on my behalf. This work would not have been possible without their constant encouragement from a very early age to pursue my passion for meteorology. In addition, the dedication that they have shown to their country through military service and to their family has given me a tenacity to set high goals for myself and to work towards achieving them. Most of all, I must thank my wife, Laura Belanger. Very few people witnessed the struggles, made the sacrifices, and provided the encouragement like she did during this process. And without her, this dissertation would not have been completed at Georgia Tech. I look forward to beginning a new chapter in our lives together now that this step is complete!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

4D-Var	Four-Dimensional Variational Assimilation
AEJ	African Easterly Jet
AEW	African Easterly Wave
AMM	Atlantic Meridional Mode
AMO	Atlantic Multidecadal Oscillation
ATCF	Automated Tropical Cyclone Forecast system
BS	Brier Score
BSS	Brier Skill Score
CAPE	Convective Available Potential Energy
CISK	Conditional Instability of the Second Kind
CFS-R	Climate Forecast System-Reanalysis
Decay-SHIPS	SHIPS with inland decay wind model
DOT	Distance of Tornado from Tropical Cyclone Center
ECMFS	ECMWF Monthly Forecast System
ECMWF	European Centre for Medium-Range Weather Forecasts
EDA	Ensemble of Data Assimilations
ENSO	El Niño-Southern Oscillation
EPAC	East Pacific Ocean
ERA-15	ECMWF Re-Analysis-15
ERA-40	ECMWF Re-Analysis-40
ERA-Interim	ECMWF Re-Analysis-Interim
ESRL	Earth System Research Laboratory
Extended Recon	Extended TC-Tornado Reconstruction since 1920
FAR	False Alarm Rate
GFDL	Geophysical Fluid Dynamics Laboratory

GFDN	U.S. Navy's version of the GFDL
GFS	Global Forecast System
HOPE	Hamburg Ocean Primitive Equation
HR	Hurricane
HWRF	Hurricane Weather Research and Forecasting
HURDAT	Hurricane Database
IMD	Indian Meteorological Department
ITCZ	Inter-Tropical Convergence Zone
JTWC	Joint Typhoon Warning Center
LGEM	Logistic Growth Equation Model
MDR	Main Development Region
MJO	Madden-Julian Oscillation
NAO	North Atlantic Oscillation
NATL	North Atlantic Ocean
NCAR	National Center for Atmospheric Research
NCDC	National Climatic Data Center
NCEP	National Center for Environmental Prediction
NCEP/NCAR	NCEP/NCAR I Reanalysis
NEMO	Nucleus for European Modeling of the Ocean
NEMOVAR	NEMO Variational Data Assimilation System
NHC	National Hurricane Center
NIO	North Indian Ocean
NOAA	National Oceanic and Atmospheric Administration
NOGAPS	Navy Operational Global Atmospheric Prediction System
OAR	Office of Oceanic and Atmospheric Research
PDF	Probability Density Function
PDO	Pacific Decadal Oscillation

POD	Probability of Detection
PSD	Physical Sciences Division
q-to-q	quantile-to-quantile
Recon	TC-Tornado Reconstruction since 1948
ROC	Relative Operating Characteristic
ROCI	Radius of Outer Closed Isobar
RPS	Ranked Probability Score
RSMC	Regional Specialized Meteorological Centre
SHIPS	Statistical Hurricane Intensity Prediction Scheme
TC	Tropical Cyclone
TD	Tropical Depression
TIGGE	THORPEX Interactive Grand Global Ensemble
UKMET	United Kingdom Meteorological Office
VarEPS	Variable Resolution Ensemble Prediction System
WISHE	Wind-Induced Surface Heat Exchange
WMO	World Meteorological Organization

SUMMARY

The spatial and temporal complexity of tropical cyclones (TCs) raises a number of scientific questions regarding their genesis, movement, intensification, and variability. In this dissertation, the principal goal is to determine the current state of predictability for each of these processes using global numerical prediction systems. The predictability findings are then used in conjunction with several new statistical calibration techniques to develop a proof-of-concept, operational forecast system for North Atlantic TCs on daily to intraseasonal time scales.

To quantify the current extent of tropical cyclone predictability, we assess probabilistic forecasts from the most advanced global numerical weather prediction system to date, the ECMWF Variable Resolution Ensemble Prediction System (VarEPS; Hamill et al. 2008, Hagedorn et al. 2012). Using a new false alarm clustering technique to maximize the utility of the VarEPS, the ensemble system is shown to provide well-calibrated probabilistic forecasts for TC genesis through a lead-time of one week and pregenesis track forecasts with similar skill compared to the VarEPS's postgenesis track forecasts. These findings provide evidence that skillful real-time TC genesis predictions may be made in the North Indian Ocean—a region that even today has limited forecast warning windows for TCs relative to other ocean basins.

To quantify the predictability of TCs on intraseasonal time scales, forecasts from the ECMWF Monthly Forecast System (ECMFS) are examined for the North Atlantic Ocean. From this assessment, dynamically based forecasts from the ECMFS provide forecast skill exceeding climatology out to weeks three and four for portions of the southern Gulf of Mexico, western Caribbean and the Main Development Region. Forecast skill in these regions is traced to the model's ability to capture correctly the

variability in deep-layer vertical wind shear as well as the relative frequency of easterly waves moving through these regions.

Following the TC predictability studies, a proof-of-concept operational forecast system for North Atlantic TCs is presented for daily to intraseasonal time scales. Findings from the predictability studies are used in conjunction with recently developed forecast calibration techniques to render the VarEPS and ECMFS forecasts more useful in an operational setting. The proposed combination of bias-calibrated regional probabilistic forecast guidance along with objectively-defined measures of confidence is a new way of providing TC forecasts on intraseasonal time scales.

On interannual time scales, the predictability of TCs is examined by considering their relationship with tropical Atlantic easterly waves. First, a set of easterly wave climatologies for the Climate Forecast System-Reanalysis, ERA-Interim, ERA-40, and NCEP/NCAR Reanalysis are developed using a new easterly wave tracking algorithm based on 700 hPa curvature relative vorticity anomalies. From the reanalysis-derived easterly wave climatologies, a moderately positive and statistically significant relationship is seen with tropical Atlantic TCs, suggesting that approximately 20–30% of the total variance in the number of TCs on interannual time scales may be explained by the frequency of easterly waves. In relation to large-scale climate modes, the Atlantic Multidecadal Oscillation (AMO) and Atlantic Meridional Mode (AMM) exhibit the strongest positive covariability with Atlantic easterly wave frequency.

Besides changes in the number of easterly waves, the intensification efficiency of easterly waves, which is the percentage of waves that induce North Atlantic TC formation, has also been evaluated. These findings offer a plausible physical explanation for the recent increase in the number of NATL TCs, as it has been concomitant with an increasing trend in both the number of tropical Atlantic easterly waves and intensification efficiency. In addition, the easterly wave–tropical cyclone pathway is likely an important mechanism governing how the AMO and AMM modulate North Atlantic TC

frequency—more so than previous thought (e.g., Thorncroft and Hodges 2001, Hopsch et al. 2007, Kossin and Vimont 2007).

The last component of this dissertation examines how the historical variability in U.S. landfalling TCs has impacted the annual TC tornado record. To reconcile the inhomogeneous, historical tornado record, two statistical tornado models, developed from a set of a priori predictors for TC tornado formation, are used to reconstruct the TC tornado climatology. Based on the evaluation period during the most reliable portion of the TC tornado record, these models possess moderate skill in forecasting the magnitude of a tornado outbreak from a Gulf landfalling TC and have high skill in forecasting the annual number of TC tornadoes. While the synthetic TC tornado record also reflects decadal scale variations in association with the AMO, a comparison of the current warm phase of the AMO with the previous warm phase period shows that the median number of tornadoes per Gulf TC landfall has significantly increased. This change likely reflects the increase in median TC size (by 35%) of Gulf landfalling TCs along with an increased frequency of large TCs at landfall.

CHAPTER 1
INTRODUCTION

Tropical cyclones (TCs) typically have lifetimes that span time scales of days to weeks, have horizontal sizes that range from 50 km to 1000 km, and can generate extreme surface winds, torrential rainfall, inland storm surge, and tornado outbreaks. Their impact on the environment may last weeks (e.g., upper-level oceanic cooling, nutrient enrichment, enhanced primary production, etc.) to even years or decades later (e.g., ecological impacts, landscape evolution, land use change, etc.). In addition, recurving tropical cyclones may interact with the extratropical jet stream to induce significant downstream responses to the large-scale atmospheric circulation—thereby modifying extratropical weather and even modulating atmospheric predictability.

Given their complexity and often extreme impact on humanity and to the natural ecosystem, this dissertation addresses understanding the extent to which tropical cyclones and their downstream impacts are predictable and how current forecast approaches can be improved. Chapter 3 examines the predictability of tropical cyclone genesis, track and intensity for North Indian Ocean TCs during 2007–2010 on time scales from the short-term through two weeks in advance. Chapter 4 provides an evaluation of dynamically-based intraseasonal tropical cyclone forecasts for the North Atlantic basin during 2008–2009. Findings from these predictability studies are used in conjunction with recently developed forecast calibration techniques to remove model bias and to adjust model spread in the operational forecast system for NATL TCs that is described in Chapter 5. In consideration of long-term tropical cyclone variations, a climatology of easterly waves for the North Atlantic and East Pacific has been produced to ultimately better understand their relationship with TCs and large-scale climate modes of variability (Chapter 6). Finally, to examine some of the meteorological impacts from TCs, in particular TC-

induced tornadoes, a statistical model has been developed that shows promise as a real-time forecast tool to predict whether a given landfalling TC will be a prolific tornado producer and has been used to reconstruct the historical TC-tornado record (Chapter 7). Before discussing the variety of findings from these projects, relevant background information is provided to place in perspective these results within the current state of relevant research.

1.1. Tropical Cyclogenesis and Intensification

A tropical cyclone is defined as a warm-core, non-frontal, low pressure system originating over tropical or sub-tropical waters and consists of organized convection and a closed cyclonic surface wind circulation (Frank 1977). Tropical cyclones include all systems designated tropical depressions, tropical storms, or hurricanes. The tropical cyclone lifecycle consists of three main phases that include tropical cyclogenesis, intensification, and decay. Tropical cyclogenesis is the process by which convective scale processes organize to produce a finite-amplitude, tropospheric-deep, warm core with a closed surface circulation and deep convection near the center. This process is thought to result from the interaction between the large-scale environment and internal convective and mesoscale dynamical processes. Ooyama (1964) and Charney and Eliassen (1964) proposed some of the earliest ideas on TC formation by attempting to connect the convective scale with the meso-alpha scale through conditional instability of the second kind (CISK). CISK represents a cooperative interaction between deep cumulus clouds and incipient vortex formation through a positive feedback mechanism that relies on moisture convergence and convective available potential energy (CAPE). However, CISK suffers from several deficiencies including most notably the incorrect characterization that moisture convergence drives vertical ascent within pre-existing tropical disturbances and that surface fluxes of sensible and latent heat are inconsequential in maintaining moist convection.

The deficiencies in CISK led Emanuel (1986) to propose another mechanism to explain how TCs develop through wind-induced surface heat exchange (WISHE). WISHE relies on the thermodynamic disequilibrium between the ocean surface and low-level atmosphere whereby a positive feedback mechanism is established between wind speed near the surface and evaporation from the underlying ocean. Upon scrutiny with modeling and observational data, several studies confirm that an established TC vortex may intensify through WISHE (Craig and Gray 1996, Molinari et al. 1998, Zhu et al. 2004). However, this process has not been universally adopted to describe TC intensification, and alternative mechanisms have been offered by Montgomery et al. (2009) and Smith et al. (2009). The main criticism toward WISHE is that this process describes axisymmetric intensification that preserves thermal wind balance, when in reality observations show that TCs are highly asymmetric during the intensification phase. Moreover, WISHE describes the intensification change via the primary circulation of a TC and offers no indication of the role the secondary circulation might play in this process. As observations indicate that intensifying tropical cyclones are accompanied by bursts of intense convection, localized buoyancy changes in association with the secondary circulation of a TC may be an important component in the intensification process.

In addition, a pre-WISHE process is necessary to explain how a surface cyclone forms since WISHE assumes the presence of a finite-amplitude surface vortex (Saunders and Montgomery 2004, Molinari et al. 2004). Hendricks et al. (2004) in a cloud-resolving simulation of Hurricane Diana (1984) showed that prior to the formation of a tropical depression (TD) strength vortex, intense plumes of convection with deep vortical cores developed within a background environment of elevated cyclonic vorticity. Through merging and axisymmetrization, these vortical hot towers generated an upscale cascade of cyclonic vorticity from the convective scale to the meso-alpha scale. Results from this study led Montgomery et al. (2006) to posit that for a tropical depression to

form, the pre-TC environment must contain sufficient CAPE and low-level cyclonic vorticity to generate vortical hot towers through deep, moist convection. This process offered to describe the formation of a finite-amplitude surface vortex assumes the cyclone is built through the bottom-up. However, an alternative viewpoint through the top-down pathway has been offered by Ritchie and Holland (1997), Simpson et al. (1997), and Bister and Emanuel (1997). The two former studies utilize potential-vorticity dynamics to indicate that when two or more mesoscale convective vortices merge, this process produces a larger mid-level cyclonic vortex. Through 'action-at-a-distance' principals via potential vorticity, the larger vortex leads to a greater Rossby penetration depth. As the system adjusts to thermal wind balance, adiabatic ascent below the mid-level cyclonic vorticity maximum enhances low-level convergence, which increases low-level cyclonic vorticity. The major deficiency in the top-down pathway for tropical depression formation is that it is not entirely clear how a cold-core, mid-level cyclonic vortex transitions into a warm-core surface vortex with maximum winds confined to the lower troposphere.

1.2. Tropical Cyclone Forecasts

1.2.1. Tropical Cyclone Formation Forecasts

Regardless of the dynamical pathway by which a tropical depression forms, the current paradigm assumes that TC formation is ultimately governed by processes that intersect at the convective-mesoscale interface. If this is true, then there may be a finite limit to the predictability of TC genesis. For instance, in homogeneous, isotropic 3D turbulence in the Kolmogorov inertial range, the kinetic energy per unit wave number, k, follows the well-known $k^{-5/3}$ power law with the eddy turnover time (i.e., the time it takes for small-scale errors to impact the large-scale) proportional to $k^{-2/3}$ (Frisch 1995). The implication is that the time it takes for uncertainty to propagate from higher to smaller wavenumbers is finite and should lead to atmospheric predictability for phenomena in

this inertial-range of only a few days (Palmer 2000). Currently, there is evidence to suggest the $k^{-5/3}$ inertial-range exists in the real atmosphere especially for phenomena whose horizontal scales are order ~100 km (Lilly 1983, Cho et al. 1999). However, for larger horizontal scales (i.e. upper meso-alpha to synoptic), energy dissipation in this inertial range is proportional to the k^{-3} power law, which is associated with the forward enstrophy cascade from large scales to small scales (McWilliams 1984). This finding at large scales agrees with the energy cascade for 2D isotropic homogeneous turbulence in which the eddy turnover time is independent of wave number and suggests that extended-range predictions are possible assuming sufficient observations at smaller scales are readily available (Charney 1971).

The connection to tropical cyclogenesis predictability is that this process may actually encompass two sets of inertial-ranges for energy and error growth given the overlap in spatial scales. During the gestation stage, where a pre-cyclonic surface vortex or mid-level cyclonic vortex is forming at meso-beta scales, error growth may follow the $k^{-5/3}$ power law. However, once a pre-existing cyclonic vortex at mid-levels or the surface has been established and the horizontal aspect ratio (horizontal to vertical) is large, then error growth may follow the k^{-3} power law. This transition may explain why recent global numerical weather prediction systems, especially those that utilize large ensembles, are increasingly able to generate skillful TC genesis forecast beyond 48 hr, once a pre-existing cyclonic vortex is present in the initial conditions (see examples below).

A variety of statistical and dynamical techniques have been developed in an attempt to forecast TC formation at larger lead-times. Using the statistical relationships between large-scale environmental variables and TC formation, Gray (1968) developed the first seasonal TC genesis parameter, which was shown by McBride and Zehr (1981) to be moderately correlated with global TC genesis regions. DeMaria et al. (2001) created an improved, statistical TC genesis model that has since been adapted into an

operational, real-time product. The operational model has forecast skill beyond climatology at predicting TC genesis, which led the National Hurricane Center (NHC) in 2008 to begin issuing real-time, albeit subjectively-determined, probabilities of TC formation out to 48 hours (Brown et al. 2008). In fact, the forecast skill in these 48-hr outlooks has led the NHC to conduct internal evaluations of the forecast utility in 120-hr genesis outlooks (Eric Blake, personal communication).

Although the ability of DeMaria et al. (2001) to predict TC genesis is noteworthy, a fundamental problem in their modeling approach is that it does not assess the probability of TC development for individual tropical disturbances (e.g., African easterly waves, equatorial waves, mesoscale convective systems, upper-level potential vorticity anomalies) by recognizing that various pathways for TC genesis exist (McTaggart-Cowan et al. 2008) and that each pathway likely has varying degrees of predictability. In addition, this model is said to assess the large-scale environment's role in TC genesis, but uses the seasonal climatology of TC formation as a primary predictor (similar to intraseasonal and seasonal TC model predictions) even though this parameter is not a physically-based measure of TC genesis potential (Emanuel and Nolan 2004, Frank and Roundy 2006, Leroy and Wheeler 2008).

In addition to statistical methods, the predictability of TC genesis has been assessed using global numerical weather prediction (NWP) systems. The improved ability of global models to predict TC genesis is remarkable considering that less than a decade ago, Beven (1999) and DeMaria et al. (2001) showed that most global models had no significant forecast skill due to high false alarm rates in predicted TC genesis. Recently, Pasch et al. (2006) verified a set of TC genesis forecasts from the GFS, UKMET, and NOGAPS models for the 2005 hurricane season. For most TCs, the GFS and UKMET models had relatively high probabilities of detection beginning three days in advance of TC formation. In addition, Fontan and Cabanes (2008) explored the predictability of TC genesis in the southwest Indian Ocean using the ECMWF Variable

Resolution Ensemble Prediction System global model. Using 35 TCs from the period 2004–2005, they showed that average TC genesis forecast probabilities were 80% two to three days in advance of TC formation and remained above 50% through seven days. The potential skill that exists in extended-range TC genesis forecasts especially using ensemble prediction systems is considered further in Chapter 3.

1.2.2. Tropical Cyclone Track and Intensity Forecasts

In the North Atlantic, tropical cyclone track forecasts have been issued routinely through 72 hours since the 1970s. In 2001, the NHC expanded these track outlooks to cover the next 120 hours. Forecast tracks at lead-times of 48 hr have undergone significant improvement since that time, as track errors averaged near 250 n mi during the 1970s and are now around 100 n mi during the present decade. The improvement at 72-hr lead-time is even more dramatic as track errors have been reduced by 60%, with most of the improvement coming during the decade 1990–1999 (Rappaport et al. 2009).

Much of this increase in forecast skill can be attributed to advancement in global NWP systems. Prior to the early 1990s, statistical or statistical-dynamical techniques generated the most skillful tropical cyclone track forecasts through 48 hr and 72 hr (DeMaria 1996). Since that time, global and regional NWP systems have produced superior track forecasts due to better data assimilation systems, increased availability of satellite observations, improved model physics, and finer spatial resolutions. Specifically, the GFS model began showing significant improvement during the period 1997–2000 due to an upgrade in its data assimilation system and the incorporation of satellite radiances from polar-orbiting satellites. In addition, NOAA began using a modernized dropwindsonde based on GPS technology around this time (Hock and Franklin 1999). The Gulfstream IV jet aircraft releases these sondes routinely to conduct synoptic surveillance of a tropical cyclone's environment especially when a system poses threat to land. This data source has led to a meaningful improvement (10–15%) in short-

term track forecasts through 48 hr, but no improvement is seen at lead-times beyond 84 hr (Rappaport et al. 2009).

During the most recent decade, global numerical weather prediction systems continue to provide the most superior track forecasts. Although considerable investment has been made into high-resolution regional modeling through the Geophysical Fluid Dynamics Laboratory (GFDL) model in the mid-2000s and in the Hurricane Weather Research and Forecasting (HWRF) model since the late 2000s, there has only been one season (2003) in which either of these high-resolution regional models provided superior guidance to their global numerical weather prediction system counterparts at extended lead-times of 120 hr. Since 2010 the GFS model, from which HWRF and GFDL derive lateral boundary conditions, has outperformed both high-resolution regional models at lead-times of 96 hr and 120 hr in the North Atlantic. However, this finding does not apply in the tropical East Pacific, as the high-resolution regional models routinely provide track forecasts that are competitive or even superior to global NWP at extended forecast lead-times. Besides the GFS, beginning in 2006 the ECMWF atmospheric deterministic model has undergone a number of important upgrades (for more information please see Section 2.2.1). In recent years, this model is generally regarded as providing the most superior track forecasts beyond 48 hr for tropical cyclones in the North Atlantic (Cangialosi and Franklin 2011).

Although tropical cyclone track forecasts continue to show steady improvement, intensity forecasting remains extremely difficult. In fact, since 1990, the accuracy of tropical cyclone intensity forecasts has shown no overall improvement but with some small improvement in forecast skill relative to climatology (Rappaport et al. 2012). The minor increase in skill is almost entirely attributed to statistically-based forecasts through the Statistical Hurricane Intensity Prediction Scheme (SHIPS), Decay-SHIPS, and the Logistic Growth Equation Model (LGEM). In fact, LGEM is now regarded as the most

skillful intensity forecast model (dynamical or statistical) especially at extended forecast lead-times in the North Atlantic (Cangialosi and Franklin 2011).

Still, the greatest challenge with intensity forecasting is anticipating periods of rapid intensification change especially prior to landfall. Recently, a statistical rapid intensification index was developed through NOAA's Joint-Hurricane Testbed that provides a 24-hr probabilistic outlook for the likelihood that a TC will undergo rapid intensification, defined as a 24-hr intensity increase of 30 kts (15.4 m s^{-1}; Kaplan et al. 2010). After evaluating the prediction system in the North Atlantic and Northeast Pacific during 2006 and 2007, the technique was shown to have a modest ability to predict rapid intensification. However, the relatively high false alarm rates for both basins demonstrate the continual difficulty in correctly anticipating rapid intensification even at rather short lead-times.

1.2.3. Intraseasonal Forecasts

In recent years, developing skillful tropical cyclone forecasts using statistical (e.g., Leroy and Wheeler 2008) and dynamical models (e.g., Bessafi and Wheeler 2006, Vitart 2009) on intraseasonal time scales has become an active research topic. Instead of providing forecasts for individual tropical cyclones, typically the approach on intraseasonal time scales is to identify regions that show elevated risk of tropical cyclone genesis and track movement (Vitart et al. 2009, Belanger et al. 2010), which may also include regional landfall risk assessments (Gall et al. 2011). However, some studies (Elsberry et al. 2010, Fu and Hsu 2011) have examined the predictability of individual tropical cyclones on these extended time scales particularly in basins that show the greatest predictability (e.g., Northwest Pacific and North Indian Oceans). For some notable tropical cyclones in recent years (e.g., Nargis and Nuri 2008), these forecasts provided a significant likelihood of TC development two to three weeks in advance. Still, the major deficiency in evaluating predictability on a per tropical cyclone basis

beyond two weeks is that on average these forecasts have a high number of false alarms with a low probability of detection.

Whereas on synoptic time scales the principal source of predictability is derived from memory of the atmosphere's initial state, on intraseasonal time scales statistical and dynamical methods have attempted to extract forecast predictability from the Madden-Julian Oscillation (MJO)—a convectively coupled mode of tropical atmospheric and oceanic variability that projects most strongly on the 30–60 day time scale (Madden and Julian 1972, Wheeler and Kiladis 1999). This intraseasonal climate mode is not only a source of extended predictability but also modulates the frequency of tropical cyclone formation in most ocean basins. Maloney and Hartmann (2000a,b) showed that when the convectively active MJO is in its westerly phase over the East Pacific and western Atlantic, TC activity across the Gulf of Mexico and western Caribbean Sea is enhanced. Maloney and Hartmann attributed this variability in TC activity to a modulation of deep-layer (850–200 hPa) westerly wind shear, whereas Camargo et al. (2009) showed that atmospheric column moistening due to the coupling of increased large-scale ascent and enhanced low-level absolute vorticity could explain the observed change in TC frequency. Suppressed TC activity is found to occur with the reverse orientation of the MJO: in its easterly phase and with convective suppression in the eastern Pacific. In other basins such as the North Indian Ocean and Northwest Pacific, tropical cyclone frequency is more strongly modulated by the phase and amplitude of the MJO. For instance, in a recent study by Krishnamohan et al. (2012), they show that around 80% of all tropical cyclones in the North Indian Ocean form when the convectively-active phase is centered in the Indian Ocean or over the Maritime continent. The proposed physical mechanism to explain the large increase in tropical cyclone activity is due to an enhancement of low-level environmental relative vorticity concomitant with a reduction in deep-layer vertical wind shear when the MJO is centered over the eastern Indian Ocean and the Maritime continent.

1.3. The Tropical Cyclone–Easterly Wave Relationship

To understand the dynamic and thermodynamic mechanisms governing the variability in tropical cyclone frequency on time scales of seasons to decades, it is important to examine how the pathways to tropical cyclone formation have evolved over the historical record. The predominant mechanism for tropical cyclone formation in the North Atlantic is the non-baroclinic pathway (McTaggart-Cowan et al. 2008). In this pathway, westward-propagating, synoptic-scale transient disturbances in the tropics, also known as easterly waves, induce approximately 60% of all hurricanes in a given season. In the East Pacific, this pathway is even more common, although the origin of these easterly waves remains a topic of considerable debate. Earlier research indicated these waves originate over Africa or the Caribbean and propagate into the East Pacific (Frank 1975, Pasch and Avila 1994, Avila and Guiney 2000, Raymond et al. 2006). Recent studies by Ferreira and Schubert (1997), Serra et al. (2008) and Toma and Webster (I,II 2010) have challenged this paradigm by providing observational and theoretical evidence to suggest many of these waves form in situ in the tropical East Pacific and are generated through barotropic or inertial instability of the transient Inter-Tropical Convergence Zone (ITCZ).

In the tropical Atlantic, the relationship between interannual variability in the number of tropical cyclones and easterly waves has only recently been examined (Thorncroft and Hodges 2001, Hopsch et al. 2007, Agudelo et al. 2011). Thorncroft and Hodges analyzed ERA–15 reanalysis data for the years 1979–1998 and found a positive correlation between the frequency of African easterly waves (AEWs) and downstream NATL TCs. Although this result was not statistically significant given the short data record, it is suggestive of a causal, dynamical relationship between the frequency of pre-existing disturbances and the seasonal number of tropical cyclones. Hopsch et al. (2007) expanded upon these findings by producing a detailed analysis of coherent vortical structures that move off West Africa using the ERA–40 reanalysis for the years 1958–

2002 and related their variability to TC formation in the eastern Main Development Region (MDR). The MDR is defined here as the region between 10^{o}N–20^{o}N, 60^{o}W–20^{o}W. They find that easterly waves with larger initial amplitude, greater vertical coherency, and deeper convection upon leaving Africa are more likely to produce tropical cyclones immediately downstream in the eastern Atlantic basin. However, their analysis did not find a statistically significant relationship between TCs and easterly waves on annual time scales, although this result was in contrast to the positive, statistically significant correlation that was found between Atlantic tropical cyclones and 2-6 day filtered meridional winds. Since the Hodges et al. (1995) tracking algorithm, which was used by Hopsch et al. (2007), is designed to track coherent vortical structures, it is unclear to what extent the westward-propagating disturbances that were tracked are a mixture of synoptic-scale easterly waves along with mesoscale convective vortices. Given the apparent disagreement in the relationship between 2-6 day filtered meridional winds, which are a proxy for easterly waves, with the vorticity tracking results, this deficiency highlights the need for a more robust tracking approach to develop a climatology of easterly waves.

Recently, Agudelo et al. (2011) conducted a diagnostic assessment of developing versus nondeveloping African easterly waves to determine the role of the large-scale environment in tropical cyclone formation. Besides identifying the environmental variables that showed the greatest probabilistic separation between developing and nondeveloping waves, they also evaluated how the intensification efficiency of African easterly waves has evolved for the period 1980–2001 using the ERA-40 reanalysis. They find that interannual variability in large-scale environmental variables such as column integrated diabatic heating, mid-level specific humidity, deep-layer wind shear and sea surface temperatures explain half of the total variance in easterly wave intensification efficiency. However, both Hopsch et al. (2007) and Agudelo et al. (2011) analyses were

based on the ERA-40, which raises the question of reanalysis sensitivity with these findings given the limitations with this reanalysis dataset (see Section 2.1.1).

To better understand the relationship between easterly waves, tropical cyclones and the large-scale environment, a robust climatology of easterly waves across the Atlantic and East Pacific is needed. To this end, in Chapter 6 a new easterly wave tracking dataset is presented along with the resulting climatology for easterly waves in the East Pacific, tropical North Atlantic and Africa. Additionally, some initial analyses relevant to the tropical cyclone–easterly wave relationship on interannual time scales are presented.

1.4. Tropical Cyclone Impacts: TC-Induced Tornado Outbreaks

An often-overlooked impact from U.S. landfalling tropical cyclones is that most TCs spawn tornadoes (Gentry 1983). Although a majority of these tornadoes are weak, there have been cases when significant death and destruction has resulted. Hurricane Ivan in 2004 generated an outbreak of 117 tornadoes that resulted in 47 injuries, seven deaths, and $96.9 million in property damage (NOAA's U.S. Storm Events Database). For the period 1970–1999, Rappaport (2000) found that TC tornadoes were responsible for 4% of all fatalities from U.S. landfalling TCs.

Most TC-induced tornadoes form in the outer rainbands typically 200 to 400 km from the TC center with some tornadoes occurring in the inner core and eye wall (McCaul 1991). The right-front quadrant of a TC is typically the most favorable area for tornado formation since this region usually has the highest Bulk Richardson Number shear and deepest convection (McCaul 1991, Verbout et al. 2007). Hence, TCs that make landfall from the Gulf of Mexico (where the right-front quadrant is exposed over land for a longer time) are more likely to produce tornadoes in the continental U.S. than Atlantic landfalling TCs that strike the U.S. coastline obliquely.

Recently, Verbout et al. (2007) provided statistical evidence indicating that TC recurvature is a better discriminating variable for a TC tornado outbreak than simply landfall location alone. Using synoptic composites of TCs that recurve versus those that do not, the authors showed that mid-latitude troughs provide additional deep-layer and low-layer vertical wind shear to recurving TCs, which favors mesocyclogenesis and tornadogenesis, respectively. Other variables that have been identified in the literature to increase the propensity of a tornado outbreak include landfalling TCs with large horizontal size and TCs with significant mid-level dry air intrusions (Hill et al. 1966, McCaul 1991, Curtis 2004).

Given the pronounced interannual variability in the number of U.S. landfalling tropical cyclones since 1900, an interesting research question is to understand to what extent has the variability in TC frequency and their structural characteristics over time impacted the annual number of TC-induced tornadoes in the U.S. In Chapter 7 this issue is addressed directly, but since the identification of a TC tornado is determined by a variety of sources, e.g., Doppler radar, damage surveys, and eyewitness accounts, the climatology of TC tornado reports has a large undercounting bias especially prior to the construction of the NEXRAD Doppler radar network in the mid-1990s. Thus, a statistical model has been developed to reconstruct the historical TC tornado record using a set of physically-based predictors. Whereas this approach provides the opportunity to understand the relationship between interannual variability in TC tornadoes and U.S. landfalling TCs, the statistical model is also shown to have skill in providing real-time forecast guidance for the tornadic potential of U.S. landfalling TCs.

1.5. Summary of Research Objectives

As a testament to the complexity of tropical cyclones, the following dissertation covers a variety of research problems. These topics range from evaluating the predictability of TC genesis, track movement, and intensity change on short time scales

to understanding their relationship with easterly wave variability and TC-induced tornadoes on interannual time scales. The list below summarizes the research goals for the work that follows.

- Evaluate the predictability of tropical cyclone formation, track movement, and intensity change. How does this predictability vary among ocean basins and by time scales? (Chapter 3, Chapter 4, Chapter 5)
- Determine the forecast skill in TC forecasts on intraseasonal time scales for the North Atlantic. Are there varying degrees of predictability based on initial conditions? To what extent does the MJO modulate this predictability? (Chapter 4, Chapter 5)
- Utilize the TC predictability studies to develop a proof-of-concept forecast system for TCs on both synoptic and intraseasonal time scales. What is the optimum way to statistically-render ensemble forecasts to provide more useful deterministic and probabilistic forecast information in the operational system? (Chapter 5)
- Develop a climatology of easterly waves for the tropical Atlantic and East Pacific. What is the relationship between historical variability in easterly waves and North Atlantic tropical cyclone activity? How do large-scale climate modes project onto the climatology of easterly waves? (Chapter 6)
- Consider how the interannual change in the number of landfalling TCs has impacted TC-tornado frequency in the U.S. What factors are the most important in determining whether a U.S. landfalling tropical cyclone will result in a large tornado outbreak? To what extent can these relationships be quantified to generate forecasts for the frequency of TC-induced tornadoes? (Chapter 7)

CHAPTER 2
DATA AND METHODOLOGY

This chapter discusses the principal data sources and methodology used throughout the following dissertation. In Section 2.1 the various global reanalysis datasets and observed hurricane databases are reviewed. In Section 2.2 background information is provided for the 15-day ECMWF VarEPS and monthly forecast systems that are used in the TC predictability studies of Chapter 3 and Chapter 4 and in the operational forecast system found in Chapter 5. To identify tropical cyclones from the forecast model fields, a variety of tropical cyclone tracking schemes have been utilized, and the configuration of each tracking algorithm is presented in Section 2.3. Finally, Section 2.4 provides a review of various deterministic and probabilistic forecast verification tools. These skill scores and other evaluation techniques are used in subsequent chapters to assess how the forecast bias and distributional errors have evolved with time in the 15-day and intraseasonal forecasts.

2.1. Datasets

2.1.1. Global Atmospheric Reanalyses

The primary observational datasets used throughout this dissertation are from atmospheric reanalyses. An atmospheric reanalysis utilizes a frozen configuration of an atmospheric numerical weather prediction system model in conjunction with a uniform data assimilation system that merges a variety of direct and indirect observations from the land surface, aircraft, rawinsondes, satellites, etc. to construct a set of global or regional state estimates of the atmosphere over an extended period of time. The primary advantage to using an atmospheric reanalysis over model-derived analysis fields is that it reduces any artificial discontinuities resulting from changes in the real-time data

assimilation system or model configuration over time. However, a reanalysis represents only one view of the atmospheric state, as there is extreme dependency on the configuration of the atmospheric model, its data assimilation system, and the set of ingested observations. For this reason, four atmospheric reanalyses are used throughout the dissertation, and where appropriate (e.g., the easterly wave climatology in Chapter 6), results from these products are compared to ensure the findings are robust. These four reanalyses include the: Climate Forecast System-Reanalysis (CFS-R), NCEP/NCAR I reanalysis, ECMWF Re-Analysis-40 (ERA-40), and ECMWF Re-Analysis-Interim (ERA-Interim).

The CFS-R is considered the most sophisticated global reanalysis to date, created using a high-resolution global model that was coupled between the atmosphere–ocean–land surface–sea ice system and covers the period 1979–2010. The CFS-R has a 50 km (T382) horizontal resolution, includes 62 vertical levels and features the same data assimilation system, Gridpoint Statistical Interpolation, that is used in NCEP's GFS model to generate initial conditions (Saha et al. 2010). In this study, the global CFS-R dataset was utilized on a geographic coordinate system using a regular latitude-longitude grid with 0.5^{o} x 0.5^{o} horizontal resolution.

In addition to the CFS-R, the NCEP/NCAR I reanalysis is also used.[1] This reanalysis was constructed from a data assimilation system and model identical to the Global Forecast System model that was operational at NCEP as of January 1995, except that it was integrated at a reduced horizontal resolution of approximately 210 km (T62) with 28 vertical levels (Kalnay et al. 1996). Although this reanalysis dataset has been used extensively in global and regional large-scale atmospheric studies, there are a number of issues that have been identified, including: spurious moisture and humidity

[1] NCEP/NCAR reanalysis data is provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, from their web site at http://www.esrl.noaa.gov/psd/.

data especially in the tropics, improper assimilation of upper tropospheric satellite radiances, and some periods of missing data during the period 1948–2010. A complete list of known problems with the NCEP/NCAR I reanalysis is available from ESRL and the CPC.[2,3]

Two European reanalysis products are also used, and they include the ERA-40 and ERA-Interim reanalyses. The ERA-40 reanalysis uses a 3D variational data assimilation technique, has a horizontal resolution of approximately 120 km (T159), and features 60 vertical levels (Uppala et al. 2005). A number of problems have been identified in the literature since its release. These problems include: the stepwise increase in analysis quality due to new satellite radiance measurements over time and excessive precipitation and humidity across tropical oceanic regions especially in the period since 1991 (Simmons et al. 2007). A complete list of these issues is available from ECMWF.[4]

The most recent European reanalysis, known as ERA-Interim, is also used. This reanalysis features the most sophisticated data assimilation system of the four reanalyses considered here as it uses a 12-hr 4D variational data assimilation system. The ERA-Interim also includes variational bias correction to better handle satellite radiance data and has a horizontal resolution of approximately 80 km (T255) with 60 vertical levels. The ERA-Interim reanalysis is produced using an older version of ECMWF's Integrated Forecast System model, which was operational from December 2006 to June 2007 (Simmons et al. 2007, Dee et al. 2011). In this study, both the global ERA-40 and ERA-Interim datasets were utilized on a geographic coordinate system using a regular latitude-longitude grid at 1.0^{o} x 1.0^{o} horizontal resolution.

[2] NOAA's ESRL: http://www.esrl.noaa.gov/psd/data/reanalysis/problems.shtml
[3] NOAA's CPC: http://www.cpc.ncep.noaa.gov/products/wesley/reanalysis.html#problem.
[4] ECMWF: http://www.ecmwf.int/research/era/Data_Services/section3.html.

2.1.2. Tropical Cyclone Datasets

The Hurricane Database (HURDAT) is used to identify tropical cyclones in the North Atlantic and the East Pacific (EPAC) and to highlight which TCs made landfall in the U.S. HURDAT is the official TC tracking dataset for the NATL and EPAC and consists of 6-hr location coordinates and maximum wind speed estimates of tropical or subtropical systems that obtained tropical storm strength or greater and is updated annually by the National Hurricane Center (Jarvinen et al. 1984, Neumann et al. 1999). For TCs in the North Indian Ocean, the Joint Typhoon Warning Center's (JTWC) best-track dataset is used (Chu et al. 2002). Information taken from the best-track datasets includes the location and time of TC genesis—defined as the first occurrence of tropical depression advisories—along with 6-hr location and intensity estimates. In addition, the location, intensity, and date of TC landfall as listed in HURDAT is also used. Since the NHC and JTWC only archive tropical systems that reached at least tropical storm strength (i.e., maximum 1-minute-sustained surface winds of at least 30 kts), tropical cyclones with peak intensities meeting tropical depression criteria are not included in the subsequent analyses. For the North Atlantic, East Pacific, and North Indian Ocean, the historical record of TCs in HURDAT and by JTWC encompasses the period 1851–2011, 1949–2011, 1945–2011. Although the North Atlantic record west of 60^{o}W is likely reliable since 1944 (Landsea et al. 2007, Mann et al. 2007), the East Pacific record is deemed reliable for the period since 1966 due to advent of meteorological satellites (Davis et al. 1984), and the North Indian Ocean record is most reliable since 1991 (Mohapatra et al. 2011).

2.2. ECMWF 15-Day and Monthly Forecast Systems

The primary forecasts models used in the analysis that follows are derived from the ECMWF. The rationale for their selection is not happenchance, but is due to their superior forecast performance that has been well documented for both the tropics and

extratropics (Hamill et al. 2008, Johnson and Swinbank 2009, Hagedorn et al. 2012). For instance, the ECMWF high-resolution deterministic routinely provides the best track forecasts for tropical cyclones in most global ocean basins and is even superior to some operational forecast centers at extended forecast lead-times (Belanger et al. 2012, Cangialosi and Franklin 2012). In addition, the ECMWF's Variable Resolution Ensemble Prediction System (hereafter, VarEPS) is generally regarded as the most well calibrated global ensemble system to date. In fact, for the same level of forecast skill, the VarEPS provides on average an additional day of forecast lead-time relative to other global ensemble systems from the Canadian Meteorological Centre, the National Center of Environmental Prediction, and the Japanese Meteorological Agency (Hagedorn et al. 2012). Since the primary goals of this dissertation are to assess the current state of tropical cyclone predictability, and then use statistical post-processing tools to make these forecasts better, a logical starting point is to utilize the forecasts from the best global numerical prediction system to date, which today just so happens to be produced by the ECMWF.

2.2.1. ECMWF Variable Resolution Ensemble Prediction System

During the period 2007–2011, the VarEPS has undergone a number of important changes that include but are not limited to increasing the horizontal and vertical resolution of the modeling system, expanding data assimilation procedures to include a greater number of satellite radiance measurements, updating model physics (including cumulus convection parameterization schemes) and changing the generation of initial and stochastic perturbations. As of 26 January 2010, the VarEPS includes the ECMWF deterministic global model run at T1279 spectral truncation (horizontal resolution ~ 16 km) with 91 vertical levels out to ten days along with 51 ensemble members (50 perturbed members + 1 control run) and at T639 spectral truncation (horizontal

resolution ~ 32 km) with 62 vertical levels. For days 11–15, the 51-member ensemble is processed at a reduced T319 spectral truncation (horizontal resolution ~ 63 km).

As of 22 June 2010, the methodology to generate initial perturbations in the VarEPS was modified to an ensemble of data assimilations (EDA) system with initial time singular vectors (Buizza et al. 2008, 2010). Previously, ensemble perturbations were constructed using both the evolved and initial singular vector approach, which captured the fastest growing errors in the first 48 hours of the model integration along with tendency errors in the model analysis (Buizza and Palmer 1995). When a tropical cyclone was present in the initial conditions, five additional initial singular vectors were computed and perturbed in a small region enclosing each TC using a diabatic, adjoint version of the ECMWF global atmospheric model at T42 spectral truncation with 62 vertical levels (Barkmeijer et al. 2001, Puri et al. 2001). However, only temperature, winds, and surface pressures were perturbed.

With the new update, initial conditions in the VarEPS are constructed using the EDA system and initial-time singular vectors. The initial-time singular vectors are retained from the previous system to ensure sufficient spread during the model integration, although their amplitude has been reduced by 10% from the previous configuration. In addition, upper-level specific humidity is now one of the model state variables that is perturbed in addition to temperatures, winds, and surface pressure. The EDA consists of ten different lower-resolution (T399L62 versus T1279L91) 4D-Var assimilations that modify observations and sea surface temperatures through stochastic perturbations that are proportional to their respective observational errors (Isaksen et al. 2010). Model error uncertainty is incorporated using the stochastically perturbed parameterization scheme, which changes the total parameterized tendency of the model's physical processes (Palmer et al. 2009). For the first time, the new EDA technique with initial singular vectors quantifies the quality of the 4D-Var deterministic analyses and provides estimates of the flow-dependent background errors in the analysis. The EDA-

initial singular vector technique is used to initialize the VarEPS, which is run twice-daily at 00UTC and 12UTC.

2.2.2. ECMWF Monthly Forecast System

The ECMFS is an extension of the Variable Resolution Ensemble Prediction System (Buizza et al. 2007), which consists of 50 perturbed members plus a control simulation of the ECMWF general circulation model that is now integrated twice-weekly on Mondays and Thursdays. For the period 2008 and 2009, the ECMFS has a horizontal resolution at approximately 50 km (T399) with 62 vertical levels for the first ten days. For days 11–32, the model integrations occur at a reduced T255 spectral truncation or a horizontal resolution of 80 km. At day 10, the atmospheric model is coupled to the Hamburg Ocean Primitive Equation (HOPE) model, which has 29 vertical levels, a zonal resolution of 1.4^{o}, and a meridional resolution that varies from 0.3^{o} near the equator to 1.4^{o} polewards of 30^{o}. For the period 2010 and 2011, the horizontal resolution of the ECMFS was upgraded to the current real-time configuration of the VarEPS listed in Section 2.2.1. In addition, as of 15 November 2011 the VarEPS and Monthly Forecast Systems now utilize the NEMO ocean model and the NEMOVAR ocean data assimilation system (Richardson 2011).

ECMWF also generates weekly hindcasts of their real-time VarEPS, which allows for operational forecast calibration. The ECMWF weekly hindcast product is a 5-member ensemble that is integrated every Thursday from the same day and time of the year for the last 18-years. One major difference between the ECMWF hindcasts and the real-time VarEPS is that the ECMWF hindcasts utilize the ERA-Interim reanalysis for model initialization (Hagedorn et al. 2012). In addition, while the real-time VarEPS has migrated to the new EDA–initial singular vector approach to generate initial conditions, the EDA system is not used for previous years with the hindcasts. Instead, the weekly hindcasts for the last 18 years use EDA-based perturbations developed from the current

year minus 14 days. Buizza et al. (2010) shows that despite using EDA perturbations from another year, the ECMWF hindcasts have similar spread–skill relationships as the real-time VarEPS. In Chapter 5, the 2010 hindcast integrations that were created one day earlier are combined with the 2011 hindcasts resulting in a total of 180 forecasts—sufficient for a robust state estimation of the model's climate.

2.3. Tropical Cyclone Tracking Schemes

A variety of tropical cyclone tracking schemes are utilized in the following chapters of the dissertation. For the predictability studies during the medium-range and intraseasonal period, two primary tracking algorithms are used as described in Holland et al. (2010) and Vitart et al. (1997), respectively. However, when both approaches were evaluated for use in the operational forecast system, several deficiencies became apparent including: too many false alarm ensemble forecast tracks using Vitart et al. (1997), and in Holland et al. (2010), the tracking thresholds are too strict resulting in a lack of ensemble forecast tracks for some observed tropical cyclones. Ultimately, a new tracking scheme was developed for the medium-range that combined elements from both tracking algorithms, reducing the number of false alarms and ensuring a robust sample of ensemble forecast tracks for an observed tropical cyclone present in the initial conditions.

2.3.1. Tropical Cyclone 15-Day Tracking Scheme

For the tropical cyclone predictability study in the North Indian Ocean (Chapter 3), tropical cyclones in the model analysis and forecast fields are isolated using a modified version of the Suzuki-Parker tracking scheme (Holland et al. 2010). To increase the intensity retrievals from the VarEPS, 10 m winds are replaced with winds averaged in the lower troposphere (10 m, 925 hPa, and 850 hPa). The tracking scheme processes each ensemble member for tropical cyclones by first identifying candidate vortices that exhibit a local minimum in mean sea level pressure. This initial set of vortices is filtered by removing systems that do not have a maximum lower tropospheric

wind speed greater than 16 kts (8.2 m s^{-1}) and an 850 hPa relative vorticity maximum greater than $1x10^{-4}$ s^{-1}. The next component of the tracking scheme confirms that each identified vortex possess a warm core as defined using the Hart (2003) phase-space method. The cyclone phase analysis quantifies the thermal structure by assessing the lower to middle troposphere's thickness gradient across the cyclone and the magnitude of the cyclone's lower troposphere and middle to upper troposphere's thermal wind. Finally, all tropical cyclone forecast tracks are required to have a lifetime of at least one day.

In the operational forecast system (Chapter 5), the modified Suzuki-Parker tracking scheme is utilized in real-time except that additional modifications are included to ensure that tropical cyclones are robustly extracted from real-time model analyses and forecast fields. First, cyclonic vortices are no longer tracked using the local minimum in mean sea level pressure but instead utilize the local maximum in 850 hPa relative vorticity. The vorticity threshold at 850 hPa is unaltered at 1×10^{-4} s^{-1} to identify the initial location of a tropical cyclone, but has been lowered to 5×10^{-5} s^{-1} for subsequent vortex tracking. Additionally, the relative vorticity at 500 hPa is checked to ensure a local maximum is present within a $3^{o} \times 3^{o}$ domain of the 850 hPa vorticity maximum. Next, a 6-hr forecast position estimate for each identified tropical cyclone is made and assumes that tropical cyclone movement over this short-window is due to advection of the mean wind in the 850–500 hPa layer in a $5^{o} \times 5^{o}$ domain encompassing the tropical cyclone's current location. This estimated forecast position defines the small domain that is evaluated at the next forecast time step to identify whether the tropical cyclone is still active. By estimating the future location of the tropical cyclone, this step eliminates the distance of movement requirement that is utilized in the original Suzuki-Parker tracking algorithm.

Finally, the ability to specify the initial location of a tropical cyclone or tropical disturbance has been included in the operational tracking scheme. In previous instances

where the Suzuki-Parker tracking approach was utilized in real-time, one deficiency that was identified concerned forecasts where a tropical depression or weak tropical storm was present but no ensemble forecast tracks from the VarEPS were generated. These cases indicate that the various dynamic and thermodynamic thresholds used in the tracking scheme are sometimes too strict for certain tropical cyclones. Therefore, the operational forecast system now utilizes an invest data file that specifies the observed location of a tropical cyclone(s) that is available via the NHC's Automated Tropical Cyclone Forecast system (ATCF; Miller et al. 1990) or may be created locally. When the invest data file is available, a separate set of ensemble forecast tracks are created by evaluating a small domain surrounding this initial location, ensuring that the local maximum 850 hPa relative vorticity exceeds 5×10^{-5} s^{-1} in this domain, and that the maximum lower tropospheric wind exceed 16 kts (8.2 m s^{-1}). The tropical cyclone's location at future time steps utilizes the same approach described previously where an estimated 6-hr location is developed using the mean environmental wind of the tropical cyclone. Technically, these ensemble forecast tracks are not true 'tropical cyclone' tracks since they do not pass any warm core requirement check. However, the set of ensemble forecast tracks that are operationally available may be used to supplement the primary set of ensemble tropical cyclone forecast tracks especially in instances where the operational tropical cyclone tracking scheme does not produce a sufficient set of ensemble forecast tracks for an active tropical cyclone.

2.3.2. Tropical Cyclone Monthly Tracking Scheme

Similar to the 15-day time scale, two separate tracking schemes are also utilized for the monthly forecast system. First, in the intraseasonal predictability study for North Atlantic TCs of Chapter 4, TCs are identified in the model analysis and forecast fields using a tracking algorithm based on Vitart et al. (1997). The scheme locates maxima of 850 hPa relative vorticity greater than 3.5×10^{-5} s^{-1} along with local minima in mean sea

level pressure. The thermal structure from 500 to 200 hPa is then determined to confirm that a warm core center ($\leq 2^o$ latitude-longitude) is in proximity with the identified TC center and that the temperature from the warm core center decreases outward by at least 0.5^oC within 8^o of latitude-longitude distance. The purpose of imposing this warm core requirement is to ensure that extratropical cyclones are not included in the monthly tracking results. Tropical cyclone intensity is determined using the maximum 850 hPa wind speed. Tropical cyclone trajectories are available at a 12-hr time step, since this is the temporal resolution of model fields from the ECMFS.

For the operational intraseasonal forecast system (Chapter 5), tropical cyclone tracks from the ECMFS are extracted using a similar version to the Suzuki-Parker tracking scheme that is utilized in 15-day predictability study for TCs in the North Indian Ocean (Chapter 3) except that the tracking and distance thresholds have been reduced to accommodate the coarser horizontal and temporal resolution of the model forecast fields. This tracking scheme is also used to isolate tropical cyclone forecast tracks within the weekly ECMWF hindcasts. First, the initial location of tropical cyclones is determined using a 5^o x 5^o domain enclosing local minima in mean sea level pressure. Maximum wind speed averaged in the lower troposphere must exceed 10 m s^{-1} (15 m s^{-1}) for cyclonic vortices located south (north) of 30^oN within this domain, and 850 hPa along with 500 hPa relative vorticity must exceed 5 x 10^{-5} s^{-1}. To ensure the identified tropical cyclones possess a warm core, the following parameters from the Hart cyclone phase space method are utilized: the storm-motion relative thickness gradient between 925 and 700 hPa across the cyclone must be less than 12.5 and the lower and upper tropospheric thermal wind for the 925–700 hPa and 700–200 hPa layers must be less than 5. Finally, all tropical cyclone forecast tracks are required to have a lifetime of at least one day.

2.4. Forecast Verification Statistics

Through the following chapters, a variety of skill scores and other verification techniques are utilized to assess how the VarEPS, ECMFS, and the calibrated forecast systems perform with increasing forecast lead-time. These verification tools include: the Brier Skill Score (BSS), hit rate, false alarm rate, relative operating characteristic (ROC), reliability diagram, and the ranked probability score (RPS).

The BSS (Equation 2.2) measures the accuracy or relative skill of a forecast over climatology by predicting whether or not an event will occur in comparison to observations and is defined as:

$$BS = \frac{1}{N}\sum_{i=1}^{N}\left(p_i - o_i\right)^2 \qquad 2.1$$

$$BSS = 1 - \frac{BS}{BS_{ref}} \qquad 2.2$$

, where N is the number of forecasts, i is the i^{th} forecast, p is the forecast probability, o is a value of 1 or 0 depending on whether the event occurred, and BS_{ref} is the Brier score of climatology (Equation 2.1; Brier 1950, Wilks 1995).

Other performance measures more suitable for probabilistic forecast systems (Mason 2004), such as ROC scores, reliability diagrams, and the ranked probability score, are also utilized for verification. The ROC is a comparison of the hit rate and false alarm rate for a set of increasing probability thresholds and is therefore an assessment of the forecast skill conditioned on the observations. The hit rate, or the probability of detection, is the proportion of all forecasts where a forecast warning was issued correctly, and the false alarm rate, or the probability of false detection, is the proportion of all forecasts where a forecast warning was issued and did not occur along with the number of correction rejections (Wilks 1995, Mason and Graham 1999). The area under the ROC curve, known as the ROC score, is a value from 0 to 1, where 0.5 indicates random chance alone (no forecast skill) and a value of 1 is for a perfect forecast system (Mason

and Graham 1999). Besides assessing the forecast performance conditioned on the observations, the reliability of the forecasts is derived by determining how frequently observations correspond with each forecast probability category. The conditional bias is shown graphically through the reliability (or attributes) diagram where perfect reliability is indicated by close proximity to the diagonal, with values residing above or below the diagonal reflective of underconfident or overconfident forecasts, respectively (Wilks 1995).

In some instances, forecast probabilities are provided for a variety of categories (e.g. TC intensity). To evaluate how well these forecasts perform, the ranked probability score is determined which measures the average difference between the distribution of forecasts and observations across all categories. The formula to calculate the RPS is listed below in Equation 2.3

$$RPS = \frac{1}{M-1}\sum_{i=1}^{M}\left[\left(\sum_{k=1}^{M} p_k\right) - \left(\sum_{k=1}^{M} o_k\right)\right]^2 \qquad 2.3$$

, where M is the number of categories, p is the probability, o is the observation (0 or 1) for the k^{th} category (Epstein 1969). Similar to the Brier score, the RPS values may range from 0 to 1, with lower values indicative of better-forecast skill. Specifically, RPS values less than $1/M$ indicate that the forecast is better than random chance alone.

CHAPTER 3
TROPICAL CYCLONE PREDICTABILTY IN THE NORTH INDIAN OCEAN ON DAILY TIME SCALES

The work presented in this chapter is in press in Weather and Forecasting and is available through early online release (Belanger et al. 2012).

3.1. Background and Motivation

Tropical cyclones in the North Indian Ocean (NIO) have a profound impact on the littoral countries of the Arabian Sea and the Bay of Bengal. The combination of a shallow coastal plain along with a thermodynamically favorable environment allow TCs to impart high surface winds, torrential rains and significant wave heights (wave setup plus storm surge) as these systems move inland. In addition, the world's highest population density coupled with low socioeconomic conditions in the region has resulted in several landfalling TCs becoming devastating natural disasters. In fact, eight of the ten deadliest TCs of all time have occurred in the Bay of Bengal and the Arabian Sea with five impacting Bangladesh and three making landfall in India (WMO-TD No. 84). Furthermore, the highest storm tide ever recorded by a TC (45 feet) occurred in the North Indian Ocean near the Meghna Estuary, Bangladesh in 1876 (WMO-TD No. 84). These occurrences highlight the need to provide regional governments and populace in the region as much advance warning as possible.

The Indian Meteorological Department (IMD), which is the World Meteorological Organization (WMO)–designated Regional Specialized Meteorological Centre (RSMC), provides the official tropical cyclone forecasts and warnings in the North Indian Ocean. As mandated by the WMO, the IMD is required to coordinate and release their forecasts daily with each member country within the North Indian Ocean. However, ultimate responsibility for forecast development and warning dissemination

lies with each country's national meteorological service. As part of the daily operational procedure when a TC is not present in the region, the IMD is required to prepare a daily tropical weather outlook, which assesses the possibility of tropical depression development in the Bay of Bengal and the Arabian Sea. Unlike the NHC's Tropical Weather Outlook, which provides the likelihood of TC genesis during the next 48 hours (Rappaport et al. 2009), the WMO gives no mandatory time constraint for this outlook product, allowing the forecast time period to be determined separately by each RSMC. In addition, this product does not provide any quantitative, probabilistic information about the potential for tropical cyclone formation or the track for the system if formation occurs.

After a TC has reached depression status, the IMD begins issuing forecast advisories, which describes the system's past movement, current location and intensity, and its future location, translation speed, wind intensity, and maximum average surface wind speed including the highest surface wind gust. The IMD does not produce storm surge forecasts, though this region has historically experienced devastating impacts from TC-induced storm surge (Webster 2008); storm surge forecasts remain the responsibility of each country's national meteorological service (WMO-TD No. 48). In addition, the WMO only requires that these forecasts cover a time horizon of three days, while most operational numerical guidance in other basins spans a five-day forecast window (Rappaport et al. 2009).

A comprehensive literature review of American Geophysical Union and American Meteorological Society journals found no articles published in the last five years that examine the predictability of NIO TCs using the latest generation of global numerical weather prediction systems. In addition, very few studies have been devoted to assessing the performance of ensemble prediction systems for tropical cyclones. Recently, Dupont et al. (2011) have assessed how well ensemble-based tropical cyclone track forecasts perform in the South Indian Ocean. In particular, they show that

calibrated probabilistic forecasts from the VarEPS perform better than climatology in assessing track uncertainty for a lead-time of three days. Results from this analysis has led RSMC La Réunion to develop their uncertainty cones using the VarEPS and to extend their track forecasts from a lead-time of three to five days.

Similar to the Dupont et al. (2011) analysis, the purpose of this chapter is to assess how well TC forecasts from the VarEPS perform in the North Indian Ocean for the period 2007–2010. In essence, forecast performance will be evaluated through an assessment of probabilistic tropical cyclone genesis forecasts, pre-genesis track and intensity forecasts. This assessment is followed by a more typical comparison of post-genesis forecast track and intensity performance of the VarEPS relative to other global modeling forecast systems. After evaluating the performance of TC genesis, track, and intensity forecasts from the VarEPS in Section 3.3, recommendations on how this model guidance may be used to produce extended-range probabilistic tropical cyclone forecasts are presented in Section 3.4.

3.2. Data and Methods

The predictability of tropical cyclones in the North Indian Ocean is evaluated using the VarEPS. During the period 2007–2010, the VarEPS has undergone a number of important changes that include increasing the horizontal and vertical resolution of the modeling system, expanding data assimilation procedures to include a greater number of satellite radiance measurements, updating model physics (including cumulus convection parameterization schemes) and changing how initial and stochastic perturbations are generated using singular vectors and an ensemble of data assimilations. More information about the model configuration of the VarEPS and its changes is found in Section 2.2.1. For the analysis in this chapter, the 15-day VarEPS 00UTC forecasts were used for the period 1 January 2007 to 31 December 2010. The VarEPS forecasts were obtained through the THORPEX Interactive Grand Global Ensemble (TIGGE) project

where the North Indian Ocean domain included the region: 0^{o}N–30^{o}N, 40^{o}E–110^{o}E and were at a horizontal resolution of 0.25^{o} x 0.25^{o}.

To isolate tropical cyclones in the VarEPS analysis and forecast fields, we use a modified version of the Suzuki-Parker tracking scheme (Holland et al. 2010). Complete details of the tracking scheme are covered in Section 2.3.1. After the tracking scheme has been implemented for each ensemble member, any tracks that originate over land are removed unless the ensemble tracks are within 300 n mi of an observed tropical cyclone. Furthermore, all ensemble forecast tracks must have a lifetime of at least one day. The scheme does have limitations with respect to exceedances in lower troposphere winds and relative vorticity which can produce unrealistic track forecasts, such as in the vicinity of 10^{o}N–17^{o}N, 40^{o}E–55^{o}E. Accordingly, all ensemble tracks from this region were eliminated unless they were in association with an observed tropical cyclone.

3.2.1. Filtering Tropical Cyclone Forecast Tracks and Determining False Alarms

After applying the tropical cyclone tracking scheme and post-processing routines to the VarEPS forecasts, a tropical cyclone filtering algorithm is used to determine which ensemble track forecasts for a particular VarEPS integration are associated with an observed tropical cyclone (Figure 3.1). First, the starting time and initial position of each ensemble track forecast, $\mathbf{E}_j(x_o, y_o, t_o)$, is compared to the initial location and time of each observed tropical cyclone, $\mathbf{O}_j(x_o, y_o, t_o)$. If the ensemble forecast track is within the spatial and temporal thresholds set by the filtering algorithm, then the ensemble forecast track is associated with an observed tropical cyclone. The spatial limit is 500 n mi at t_o with an increase of 100 n mi for each 24 hour increase in forecast lead-time. The temporal limit is ± 5 days at t_o with an increase of one day for each 24 hour increase in forecast lead-time. For ensemble track forecasts within the limits, these forecasts are further subdivided into pre-genesis forecast tracks, $\mathbf{E}_{PRE\text{-}TC\,j}(\mathbf{x})$, and post-genesis forecast tracks, $\mathbf{E}_{POST\text{-}TC\,j}(\mathbf{x})$. If the ensemble forecast track precedes the date of observed tropical

cyclone genesis, defined as the date when the JTWC issues the first tropical depression advisory, then the forecast track is classified as a pre-genesis forecast track, $\mathbf{E}_{PRE\text{-}TC\,j}(\mathbf{x})$. Post-genesis forecasts tracks, $\mathbf{E}_{POST\text{-}TC\,j}(\mathbf{x})$, begin with the first tropical depression advisory and continue until the last tropical advisory has been issued for an observed tropical cyclone.

After determining which ensemble forecast tracks are associated with an observed tropical cyclone, there is an implication that all other ensemble forecast tracks must be false alarms. However, to take advantage of the probabilistic framework of the VarEPS, we have developed a false-alarm clustering algorithm such that a false alarm occurs for a particular model integration when a cluster of VarEPS ensembles produces a localized set (in space and time) of forecast tracks. The method uses the spatial and temporal thresholds from the tropical cyclone filtering routine and "*k*–means clustering" (Mirkin 1996). *K*-means clustering is an algorithm designed to divide a set of points (i.e. initial ensemble forecast coordinates) into *k*–clusters whose membership is based on the distance between each ensemble's forecast point and each cluster's mean, or centroid, location.

First, a false alarm cluster occurs when the normalized number of ensemble tracks within a cluster is in excess of this false alarm ensemble probability, P_F. The ensemble probability threshold, P_F, is pre-defined by the user and should be determined based on the end-user's needs (i.e. for few false alarms, select a high P_F). Next, all ensemble tracks, $\mathbf{E}_j(\mathbf{x})$, for a particular model integration are compared with the set of ensemble forecast tracks that were defined as either pre-genesis, $\mathbf{E}_{PRE\text{-}TC\,j}(\mathbf{x})$, or post-genesis, $\mathbf{E}_{POST\text{-}TC\,j}(\mathbf{x})$, tropical cyclone forecast tracks. Those ensemble forecast tracks that are not in the set of tropical cyclone forecast tracks become candidate false alarms, $\mathbf{E}_{CFA\,j}(\mathbf{x})$, creating a new set of ensemble forecast tracks for further analysis, i.e. $\mathbf{E}_j(\mathbf{x}) \rightarrow \mathbf{E}_{CFA\,j}(\mathbf{x})$. Thereafter, the clustering routine begins first with the assumption that all candidate false alarm tracks belong to the same cluster, i.e. $k = 1$, so that an ensemble mean initial

location $\mathbf{F}_k(\boldsymbol{x}_0)$ and forecast time $\mathbf{F}_k(t_0)$ is defined. This coordinate information then becomes the "clustering point" that the false alarm filtering routine uses in conjunction with the spatial and temporal thresholds defined previously to determine which ensemble forecast tracks are close to the cluster's ensemble mean starting time and initial location. For ensemble forecast tracks that are not within the distance and time thresholds of the cluster, these tracks are removed from further false alarm consideration unless the number of clusters changes.

After the subset of false alarm ensemble forecasts for the cluster is identified (i.e., $\mathbf{E}_{CFA\,j}(\mathbf{x}) \rightarrow \mathbf{E}_{FAC\,j}(\mathbf{x})$), then if the total number of forecast tracks within this set is in excess of the false alarm ensemble threshold P_F, then the process described previously is repeated, except that the number of clusters is increased by one ($k = 1 \rightarrow k = 2$), so that two k-means clusters are created from the initial set of ensemble forecast tracks, $\mathbf{E}_{CFA\,j}(\mathbf{x})$. Each cluster's mean coordinate information is then used in the false alarm filtering routine to identify which ensemble forecast tracks are within the cluster's starting time and initial location. From there on, if ensemble track membership of both clusters is in excess of the false alarm probability threshold, then the complete process is repeated except that three clusters ($k = 3$) are defined. This iterative process continues until the number of ensemble tracks for any cluster drops below the false alarm ensemble probability threshold. When this occurs, the finalized number of false alarm clusters for a particular model integration is then $k–1$ clusters.

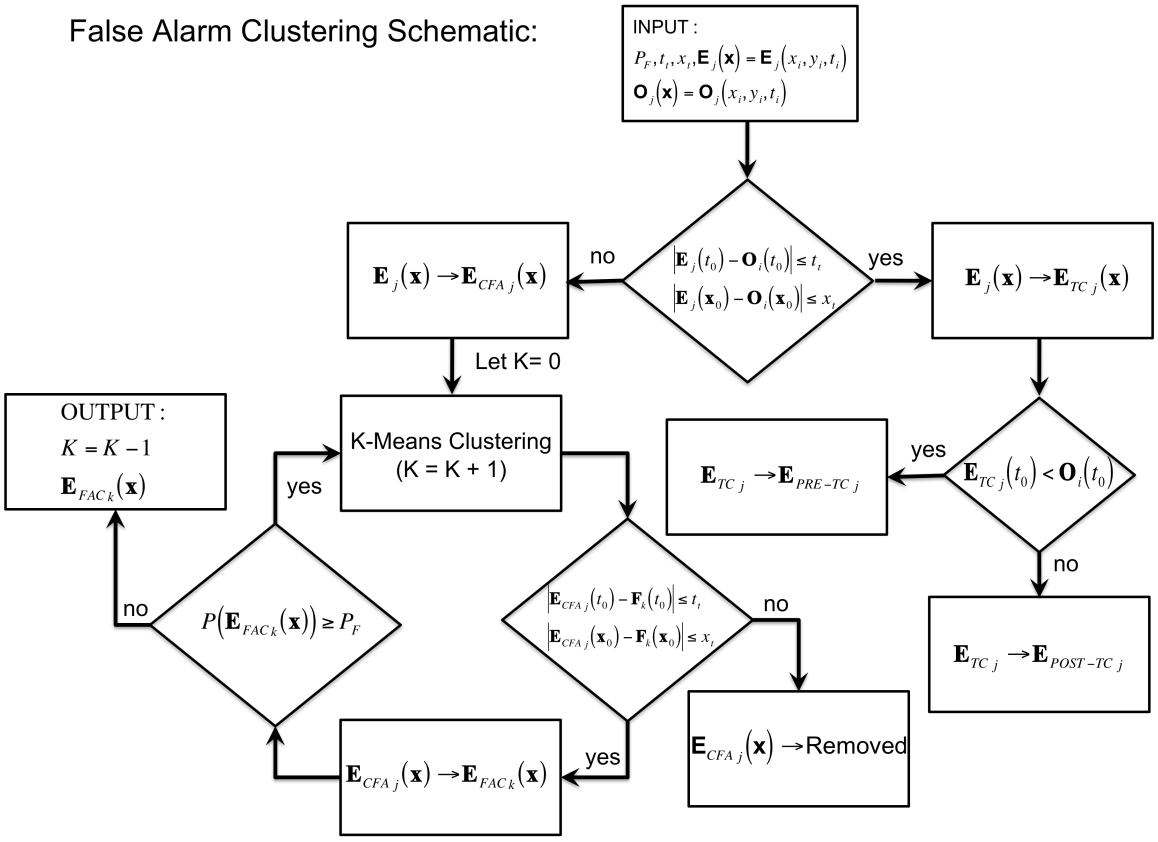

Figure 3.1: Schematic of the false alarm clustering routine. Input variables include P_F, t_t, x_t, $\mathbf{O}_i(\mathbf{x})$, and $\mathbf{E}_j(\mathbf{x})$. P_F is the false alarm probability threshold, t_t is a time threshold of 120 hr, $x_t = \left(\frac{100}{24}\right)t + 500$ (n mi), $\mathbf{O}_i(\mathbf{x})$ contains location and time information for the i^{th} observed TC during a particular season, $\mathbf{E}_j(\mathbf{x})$ is the ECMWF forecast track for the j^{th} ensemble member. Other variables include $\mathbf{E}_{TC\,j}(\mathbf{x})$ which is the ECMWF forecast track for the j^{th} ensemble member that corresponds to an observed TC. $\mathbf{E}_{TC\,j}(\mathbf{x})$ is made up of pre-genesis TC forecast tracks, $\mathbf{E}_{PRE\text{-}TC\,j}(\mathbf{x})$ and post-genesis TC forecast tracks, $\mathbf{E}_{POST\text{-}TC\,j}(\mathbf{x})$. K is the number of clusters employed in the k–means clustering algorithm, $\mathbf{F}_k(\mathbf{x})$ contains the k^{th} false-alarm cluster's mean location and starting time information. $\mathbf{E}_{CFA\,j}(\mathbf{x})$ is the 'candidate false-alarm' for each j^{th} ECMWF forecast track, $\mathbf{E}_{FAC\,k}(\mathbf{x})$ is the final false-alarm grouping for the k^{th} cluster.

3.3. Results and Discussion

In Section 3.3.1 we present an analysis of the VarEPS predictions of tropical cyclone formation in the North Indian Ocean for the period 2007–2010. The analysis is accomplished by evaluating the probability of detection and the false alarm rate as a function of forecast lead-time and increasing probability threshold. We also include an

evaluation of how well the pre-genesis tropical cyclone forecast tracks and intensities from the VarEPS perform relative to observations. In Section 3.3.2, the forecast skill for track and intensity forecasts post-genesis is evaluated. Finally, in Section 3.3.3 we examine the regional predictability of TC activity in the North Indian Ocean by evaluating the VarEPS forecasts separately for the Arabian Sea and the Bay of Bengal.

3.3.1. Tropical Cyclone Formation and Pre-genesis Forecasts

Figure 3.2a shows the spatial distribution of tropical cyclone forecast tracks for Severe Cyclone Nargis from the VarEPS initialized on 23 April 2008 00UTC, which is approximately four days prior to the initiation of tropical depression advisories by the JTWC.

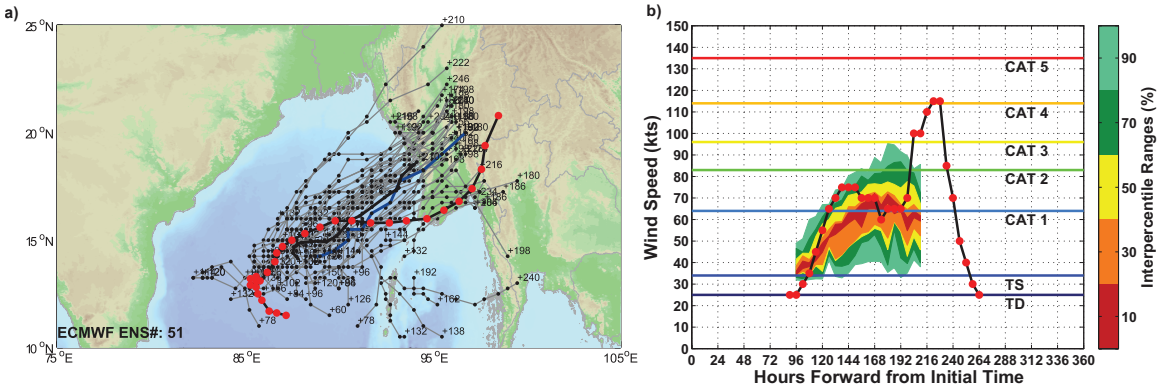

Figure 3.2: a) Example of the VarEPS forecasts for Severe Cyclone Nargis on 23 April 2008 00UTC, which was 3.75 days prior to tropical cyclone genesis according to the JTWC's best track dataset. The black line with red dots denotes the observed track of Nargis. The thin grey lines indicate unique ensemble track forecasts from the VarEPS with the thick black (blue) line denoting the VarEPS ensemble mean (control) track. b) The intensity forecast from the VarEPS is shown as a probabilistic time series in which color shading denotes percentile intervals of the VarEPS forecasts ranging from 10 to 90%. Note: The black line with red dots denotes the observed intensity of Nargis.

This case illustrates some of the forecast information that is contained within the tropical cyclone VarEPS forecasts. The VarEPS forecasts are in good agreement that the pre-tropical vortex that would become Nargis would reach advisory criteria around 27 April 2008 in the central Bay of Bengal. Thereafter, based on forecasts on 23 April, the

tropical cyclone is forecast to move generally towards the east-northeast on a track that would cause the system to make landfall in Myanmar around 30 April or 1 May with a high (60%+) probability as a hurricane and a much lower (5%) probability as a severe cyclone/major hurricane (Figure 3.2b). Nargis was observed to intensify to Category 4 hurricane level on the Saffir-Simpson intensity scale with maximum sustained winds of 115 kts. The system made landfall in southern Myanmar on 2 May 12UTC, propagating eastward across the Irrawaddy delta (Webster 2008). The forecast performance of the VarEPS for Severe Cyclone Nargis is exceptional in nearly all facets of prediction: tropical cyclone genesis was forecast correctly eight days in advance, ensemble mean track errors never exceeded 375 n mi, even based on pre-genesis track forecasts, and intensity forecasts, although underestimated, indicated a moderate (30%+) probability of Nargis reaching hurricane intensity nearly six days in advance of TC formation.

A more systematic study of all tropical cyclones in the NIO is now conducted to determine whether the Nargis forecast performance is characteristic of the VarEPS. Figure 3.3 is a ROC evaluation of the VarEPS forecasts for NIO tropical cyclone genesis using the metrics of probability of detection (POD; also known as hit rate) and false alarm rate (FAR; also known as probability of false detection) for all 23 tropical cyclones occurring during the 2007–2010 period using the false alarm clustering methodology from Section 3.2.1. The POD is a measure of the fraction of observed tropical cyclones where TC genesis was forecast correctly relative to all observed TCs. The FAR is the proportion of all forecasts where a forecast of TC genesis was issued and did not occur along with the number of correct rejections (i.e. TC genesis was not forecast to occur and it did not occur). The false alarm rate should not be confused with the false alarm ratio, which is the proportion of all forecasts where TC formation is forecast to occur but did not (Barnes et al. 2009). Since the false alarm rate is a function not only of the number of false alarms but also the number of correct rejections, the number of VarEPS forecasts included in this evaluation will modulate the false alarm rate mainly through the number

of correct rejections. If all VarEPS forecasts from 2007–2010 are included, the false alarm rate as a function of forecast probability threshold is significantly lower than if only the months traditionally associated with NIO TC activity are included. Therefore, for this analysis, the VarEPS evaluation is restricted to the months of April to June and August to December for the 2007–2010 period.

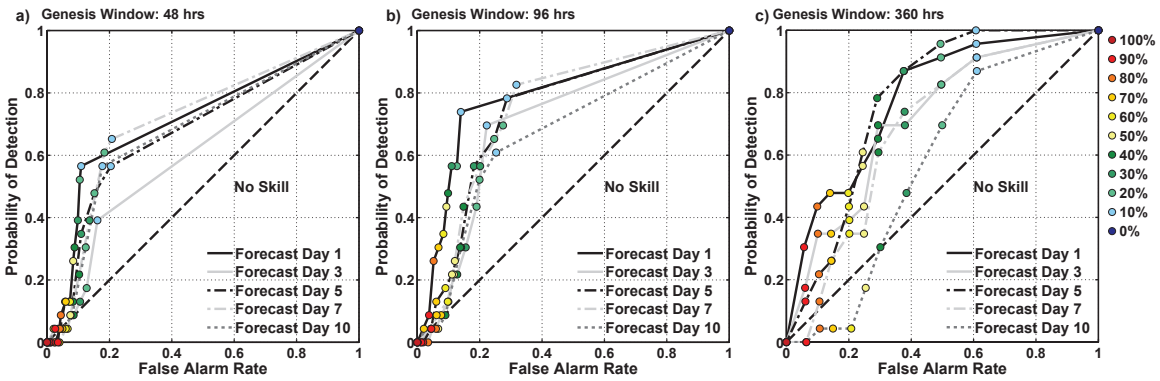

Figure 3.3: Relative operating characteristic of the VarEPS forecasts for tropical cyclone genesis during the period 2007–2010. The color-coded dots correspond to VarEPS tropical cyclone genesis probabilities (%) ranging from 0 to 100%. a) ROC is determined as a function of forecast days in advance using a 48-hr window on the date of TC genesis. b) Similar to a), except for a 96-hr window on the date of TC genesis. c) Similar to a), except the full 360-hr forecast period is used for the TC genesis evaluation.

The POD and FAR statistics are also sensitive to the time window that is used for verification. For instance, if one were to verify all one-day lead-time TC genesis forecasts, but did not require that the VarEPS correctly forecast the actual date and time of TC genesis (i.e. a 360-hr time window is used), the POD would be much larger and the FAR much smaller than if a short window centered on the time of TC genesis is required. Figure 3.3a-c shows the POD and FAR of the VarEPS forecasts using 48-hr, 96-hr, and 360-hr time windows calculated at a 10% forecast probability interval ranging from 0 to 100%. The time window is defined with respect to the time of TC genesis (e.g., a 48-hr window is ± 24-hr from t_o). Using the 48-hr time window, the VarEPS forecasts for lead-times of one to five days in advance exhibit moderate probabilities of detection (0.4 to 0.7) with very low false alarm rates (0.1–0.2) mainly for forecast probability thresholds

of 10–40% (Figure 3.3a). Although the false alarm rate never exceeds 0.2 even at extended lead-times, VarEPS forecasts made over seven days in advance tend to have low POD once forecast probability thresholds surpass 20%. This observation reflects a decreased frequency of VarEPS forecasts at long lead-times where the forecast probability exceeds 20%. If a 96-hr time window is used, the probability of detection of the VarEPS on average increases by about 0.2 and the false alarm rate by about 0.1 for forecasts of TC genesis made less than seven days in advance (Figure 3.3b). Forecasts greater than seven days in advance benefit even more from the increase time window, as the POD for a ten-day lead-time forecast increases on average by about 0.15, and the FAR increases by only 0.05 for a forecast probability threshold of 10–30%. Finally, using the full 15-day period of the VarEPS to define the time window of TC genesis, the average POD exceeds 0.7 and the FAR ranges from 0.3 to 0.6 based on forecast probability thresholds of 10–40% (Figure 3.3c).

A key component of any operational forecast system for TC genesis is knowing what combination of forecast lead-time and time window maximizes the probability of detection while minimizing the false alarm rate. If the ROC score is calculated as a function of forecast lead-time for various time windows, the optimum time window is found by determining at which time window the ROC score obtains a maximum value. For TC genesis forecasts with lead-times of one to five days in advance, a 48-hr time window around the forecast date of genesis maximizes the probability of detection while minimizing the false alarm rate. For TC genesis forecasts beyond a lead-time of five days, the ROC score is maximized if the full 15-days of the VarEPS integration is used to determine the forecast time of TC genesis.

The following analysis allows a greater understanding of the NIO false alarms that occur in the VarEPS. The spatial distribution of each false alarm cluster's ensemble starting location from 2007 to 2010 is shown in Figure 3.4a. Although the false alarms in the Arabian Sea and the southern Bay of Bengal tend to be distributed uniformly during

the 2007–2010 period, there is a relatively high concentration of false alarm clusters that stretch from the northwestern Bay of Bengal into the extreme northeastern Arabian Sea. The false alarm clusters in the northwestern Bay of Bengal tend to occur in a localized region where there is a large gradient in topography between the Bay of Bengal and the northern portion of the Eastern Ghats mountains. This concentration of false alarms may be defining the preferential track of pre-existing cyclonic vortices that move through the Bay of Bengal and into India but never become TCs. The localized nature of this false alarm concentration may indicate a relationship between current convective parameterization schemes, topography, and the forecast frequency of TC genesis in the VarEPS.

Figure 3.4b shows the false alarm ratio for the period 2007–2010 with Figure 3.4c indicating the cumulative distribution function of all false alarms as a function of forecast lead-time. The false alarm ratio obtains a peak value around 0.5 for a forecast probability level of 10% and decays nearly exponentially as the forecast probability threshold increases. At a forecast probability threshold of 25%, the false alarm ratio and the forecast probability threshold are equivalent. Figure 3.4c shows that around half of all false alarms during the 2007–2010 period occur at a forecast lead-time of 120 hr to 240 hr.

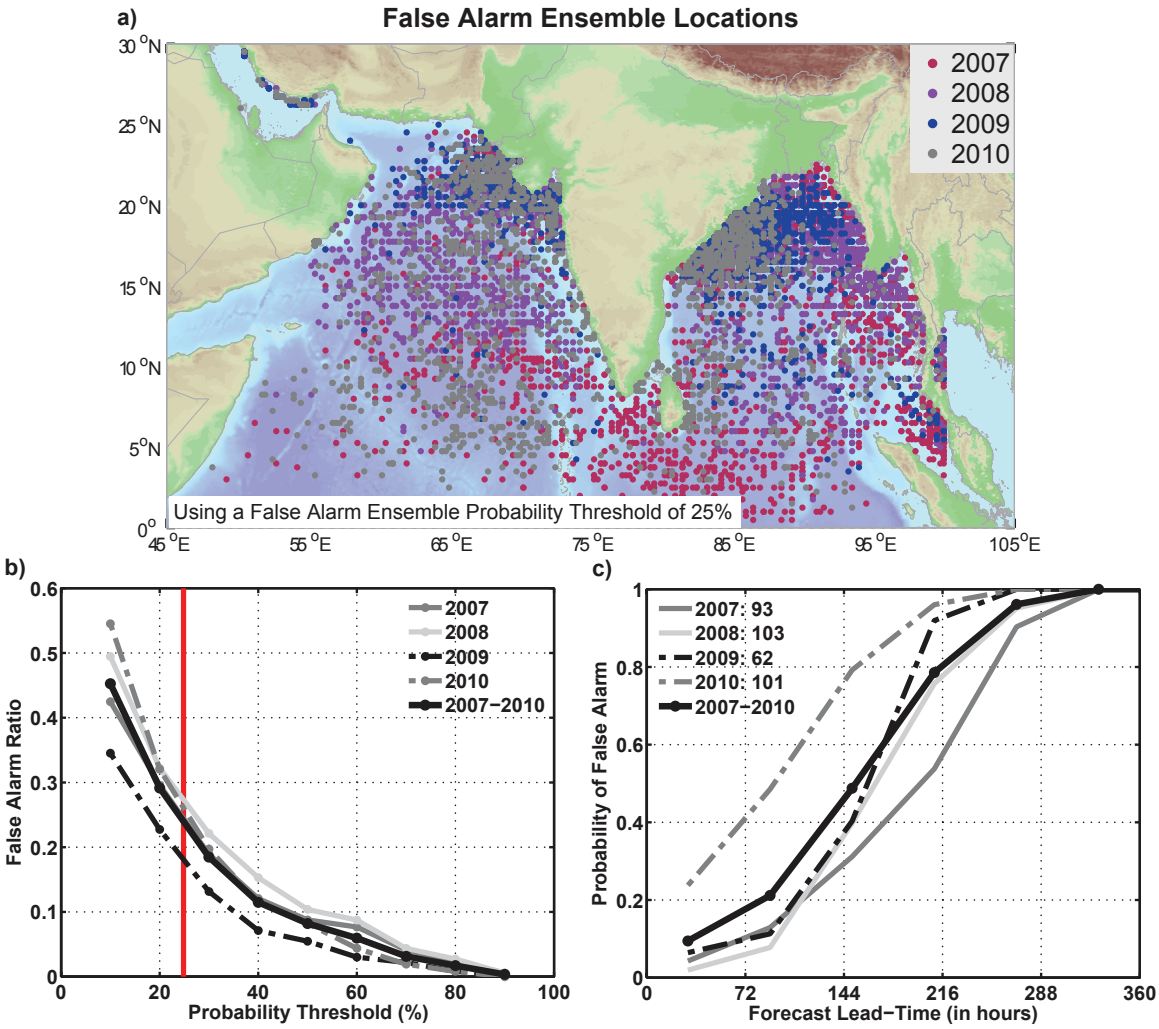

Figure 3.4: a) Initial ensemble location of each false alarm cluster from 2007 to 2010 using a false alarm ensemble threshold of 25%. b) False alarm ratio as a function of increasing false alarm ensemble probability threshold from 2007 to 2010. The red line indicates the location where the false alarm ratio is equivalent to the probability threshold. c) Cumulative distribution functions of the false alarm initiation time as a function of forecast lead-time in hours using a false alarm ensemble threshold of 25%. The legend in Figure 3.4c provides the number of false alarm clusters that occurred each year.

We now evaluate how well the VarEPS forecasts perform for track and intensity prior to TC genesis. The average ensemble mean track error at a lead-time of 24 hr is 82 n mi with a 50% interval of 51–106 n mi, and at 120 hr it is 224 n mi with a 50% interval of 116–292 n mi for all pre-genesis VarEPS forecasts during the period 2007–2010

(Figure 3.5a). From a lead-time of 24 hr to 240 hr, the mean ensemble track error growth is nearly linear at 41 n mi per day, so that by a lead-time of 240 hr, the total mean ensemble error is 409 n mi with a 50% percent interval of 186–498 n mi. In addition, the ensemble track error distribution becomes increasingly non-Gaussian as forecast lead-times increase. The implication is that beyond 72 hr, the mean ensemble track error grows larger than the maximum likelihood of the pre-genesis track error distribution. Figure 3.5a also shows that the VarEPS track forecasts perform similarly regardless of year, indicating that even though the VarEPS has undergone several major changes during the 2007–2010 period, there has not been a substantial change in forecast track performance in the NIO. To place these pre-genesis track errors in perspective, the IMD's post-genesis track errors at a lead-time of 48 hr and 72 hr typically average 162 n mi and 270 n mi, respectively. This is equivalent to VarEPS ensemble mean pre-genesis forecasts at lead-times of 90 hr and 138 hr, respectively.

Figure 3.5b shows the absolute intensity error for all pre-genesis ensemble forecasts during the period 2007–2010. Forecasts with a lead-time of 24 hr have a mean absolute intensity error of 11 kts (6 m s^{-1}) with a 50% percent interval of 6–15 kts (3–8 m s^{-1}), and by 120 hr the mean absolute intensity error grows to 23 kts (12 m s^{-1}) with a 50% interval of 7–27 kts (4–15 m s^{-1}). However, unlike the pre-genesis TC track forecasts where the 95th percentile interval is approximately 2.5 times as large as the mean ensemble track error at 120 hr, the 95th percentile interval for absolute intensity error is 3.4 times as large, reflecting a substantial negative intensity bias for several of the most intense tropical cyclones during the 2007–2010 period (not shown). In terms of the interannual variation in forecast performance of intensity, Figure 3.5b indicates that the VarEPS forecasts for 2008–2010 have on average performed substantially better than the VarEPS forecasts from 2007, with 2008 and 2009 showing three times the improvement relative to 2007. This marked change in intensity forecast skill may be due in part to the horizontal and vertical resolution increase that occurred after 2007 in the VarEPS.

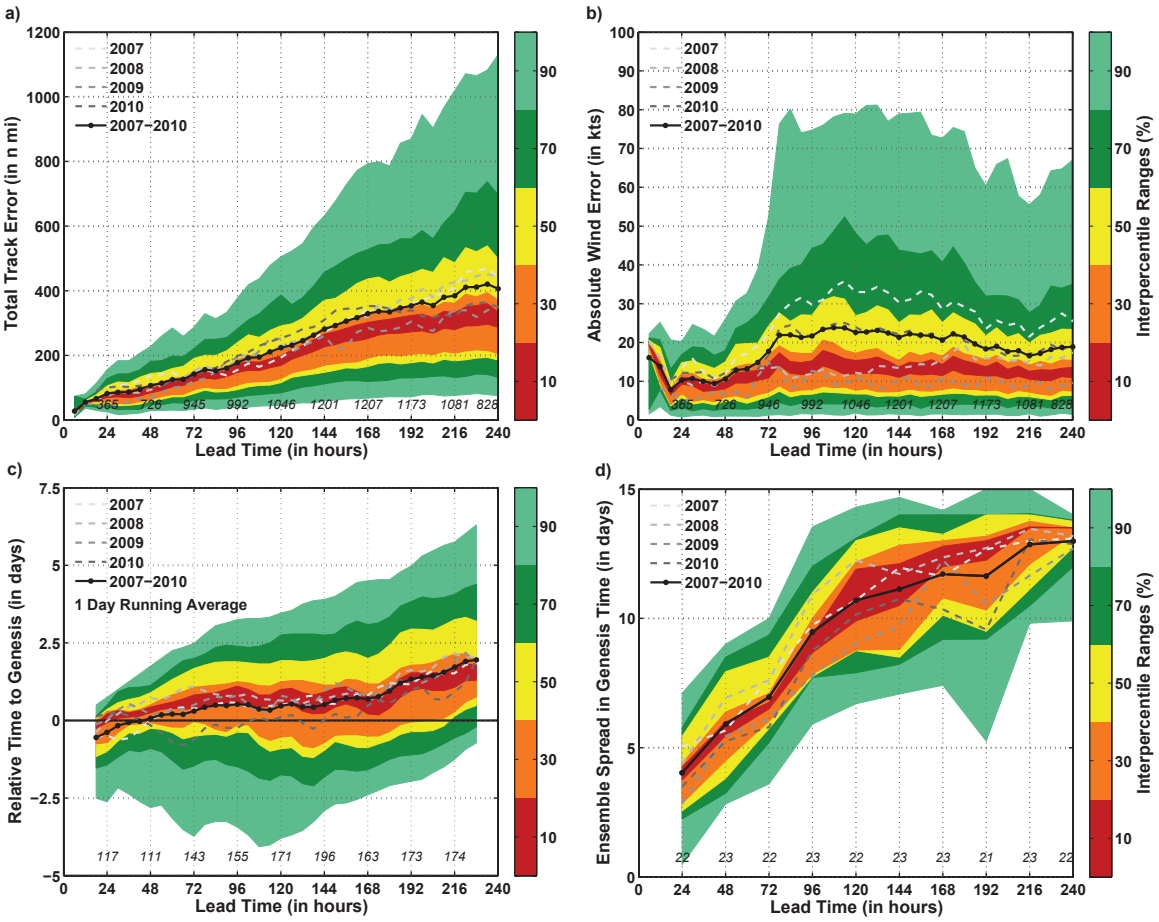

Figure 3.5: a) VarEPS pre-genesis track errors (in n mi) and b) absolute intensity errors (in kts) for all ensemble forecasts during 2007–2010. c) VarEPS forecasts of the difference between forecast time of TC genesis and observations (in days) with values greater (less) than zero indicating the VarEPS forecasts are sooner (later) than observations. Note: Values have been filtered using a 1-day running mean. d) Ensemble spread in the forecast time of TC genesis. Color shading indicates the percentile ranges for the VarEPS forecasts and the black line is the VarEPS ensemble mean. The total number of ensemble forecasts included in the verification is listed above the abscissa.

Forecasting tropical cyclone formation requires an estimate of not only the likely location of TC genesis, but also the time when a system is likely to reach advisory thresholds. Figure 3.5c shows the relative error (in days) in the VarEPS's forecast timing of TC genesis. Positive values indicate the VarEPS TC genesis forecasts occur sooner than observations, while negative values indicate a later genesis date. For the first 120 hr, the ensemble spread becomes more dispersive as forecast lead-time increases, such that at a lead-time of 24 hr, the VarEPS mean ensemble error is -0.5 days with a 50% ensemble

interval of -1 to 0.5 days. By 120 hr, the VarEPS mean ensemble error is 0.5 days but with a 50% ensemble interval of -0.9 to 1.9 days. After a lead-time of 168 hr, the VarEPS mean ensemble bias begins to increase more rapidly but with little change in ensemble spread, such that at 240 hr the VarEPS mean ensemble bias grows to 2.2 days with a 50% interval of 0.8 to 3.3 days. Although the ensemble spread is several times larger than the VarEPS mean error, the systematic growth in the mean bias as a function of forecast lead-time is a robust feature of Figure 3.5c. If one compares the mean error for the first 72 hr (-0.2 days) relative to 168–240 hr (1.3 days), this difference is statistically significant at the 99% confidence level using a bootstrap resampling test.

One possible explanation of why the VarEPS's forecasts for timing of TC genesis are well-constrained through a lead-time of 168 hr is the dispersion or spread among the VarEPS's ensembles. Figure 3.5d shows the distribution of ensemble spread in genesis time for each TC from 2007 to 2010 as a function of forecast lead-time. The ensemble spread is calculated as the difference in time (in days) of TC genesis between the earliest and latest ensemble member. As shown in Figure 3.5d, the ensemble spread in TC genesis time grows rapidly for the first seven days of forecast lead-time, then begins to grow less rapidly reaching a peak spread around 12 days after a forecast lead-time of 168 hr. While the 15-day integration period has the effect of artificially limiting the spread, the limited TC sample size from 2007–2010 makes it impossible to evaluate the statistical significance of the bias (c.f. Figure 3.5c).

3.3.2. Post-genesis Tropical Cyclone Forecasts

Figure 3.6a shows the error distribution of all VarEPS track forecasts during the 2007–2010 period indicating how the total track error statistic varies as a function of lead-time after TC genesis has occurred. From Figure 3.6a it is seen that at a lead-time of 24 hr, the mean track error is 71 n mi with a 50% interval of 35–98 n mi. From 24 hr to 120 hr, the VarEPS track errors increase linearly at a rate of 58 n mi day^{-1} such that by a

lead-time of 120 hr, the mean track error is 325 n mi with a 50% interval of 148–427 n mi. When compared to the pre-genesis track forecasts, the post-genesis track errors are larger by 15% to 30% depending on the lead-time and quantile considered of the track error distribution. Although a topic of future work, this finding suggests that the VarEPS post-genesis procedure of increasing ensemble spread through initial singular vectors in a sub-region enclosing a TC leads to greater track forecast degradation through a lead-time of 120 hr than if the procedure was not used. Similar to the VarEPS pre-genesis track forecasts, the VarEPS post-genesis track forecasts show no significant improvement in annual performance for the period 2007–2010.

To place the VarEPS's post-genesis track forecasts in perspective, Figure 3.6b compares the VarEPS control and ensemble mean forecasts with other forecasting agencies including the JTWC, the U.S. Navy's version of the GFDL (GFDN), the UKMET, the GFS, and the NOGAPS models. Since these forecasts were obtained through the U.S Navy's Automated Tropical Cyclone Forecasting system, most of the forecast guidance is limited to 72 hr, similar to the temporal limit of the JTWC's forecasts prior to 2010. Although the VarEPS control and ensemble mean forecasts on average begin with the largest initial track error, at 12 hr and beyond the VarEPS control and ensemble mean forecasts exhibit the lowest track errors among all other models. In addition, the VarEPS control and ensemble mean on average exhibit slightly lower track errors than the JTWC through a lead-time of 72 hr, although this difference is not statistically significant at the 95% confidence level. Relative to the next best performing forecast model, the VarEPS ensemble mean's 24-hr, 48-hr, and 72-hr track forecast error is on average 10%, 19%, and 27% smaller than NCEP's GFS, respectively. Since the linear track error growth per day for the VarEPS ensemble mean (41 n mi day^{-1}) is considerably smaller than the GFS (66 n mi day^{-1}), greater track forecast utility is obtained at longer lead-times with the VarEPS in comparison to other model forecasts.

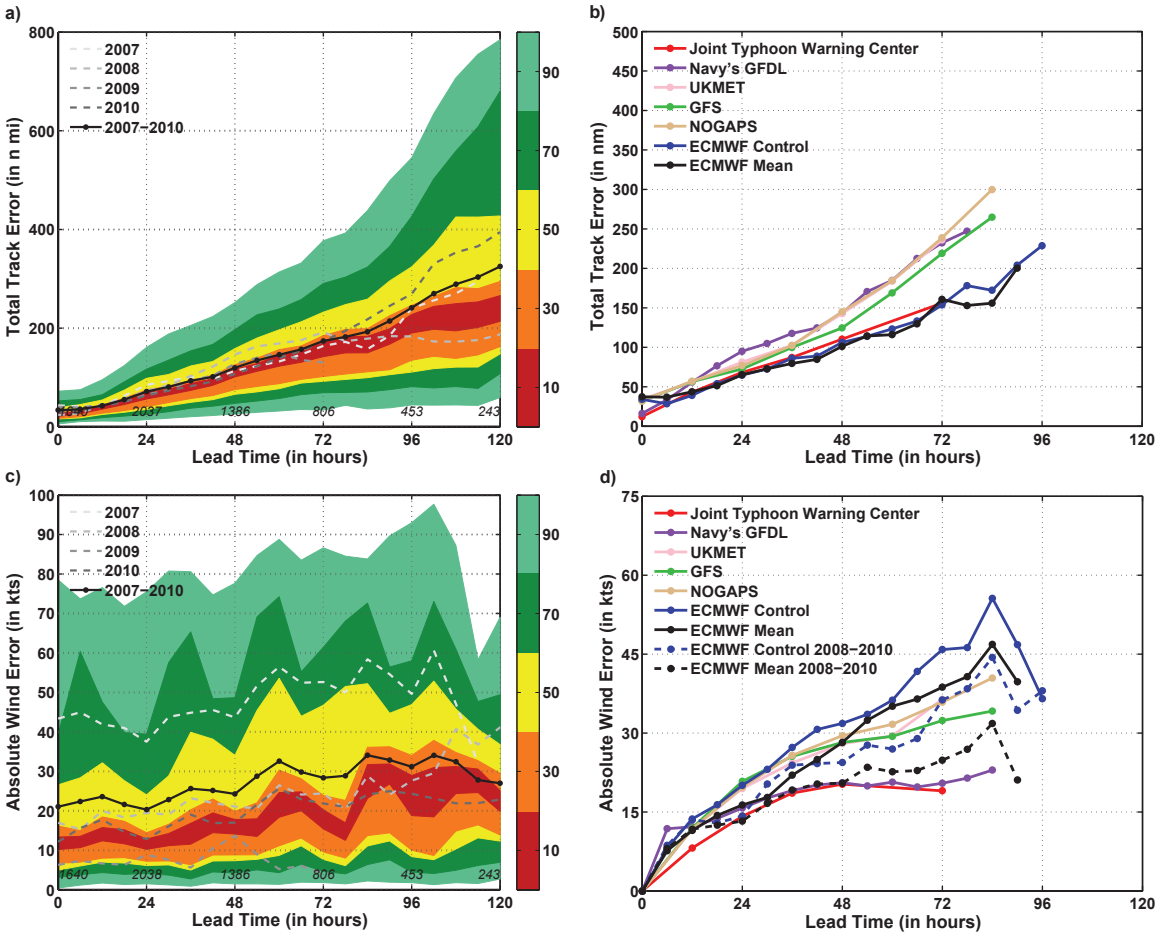

Figure 3.6: a) VarEPS post-genesis track errors (in n mi) and c) absolute intensity errors (in kts) for all ensemble forecasts during 2007–2010. Color shading indicates the percentile ranges for the VarEPS forecasts and the black line is the VarEPS ensemble mean. The total number of ensemble forecasts included in the verification is listed above the abscissa. b) Comparison of ECMWF control and ensemble mean track errors (in n mi) and d) absolute intensity errors (in kts) to other global weather models and the Joint Typhoon Warning Center for the period 2007–2010. Initial intensity bias has been removed from each model and the JTWC in Figure 3.6d.

Figure 3.6c shows the distribution of absolute intensity error for all VarEPS post-genesis forecasts during the 2007–2010 period. At the analysis time step (0 hr), the mean absolute intensity error of all VarEPS forecasts is high at 21 kts (11 m s^{-1}) with a 50% interval of 5–27 kts (3–14 m s^{-1}). By a lead-time of 72 hr, the mean absolute intensity error reaches 28 kts (15 m s^{-1}) with a 50% interval of 8–47 kts (4–25 m s^{-1}), and by 108 hr, the VarEPS mean forecasts reach their maximum intensity error of 32 kts (18 m s^{-1})

with a 50% interval of 10–46 kts (6–25 m s^{-1}). Unlike the intensity error statistics for the pre-genesis VarEPS forecasts, the interannual variation of post-genesis intensity forecasts reflects a more substantial improvement for the 2008–2010 period compared to 2007. Using the first 72 hr of lead-time as a reference, an average improvement of 67% relative to 2007 is evident.

Figure 3.6d compares the VarEPS control and ensemble mean absolute intensity error to other forecast models and the JTWC after removing the initial intensity bias. Generally for the 2007–2010 period, the VarEPS begins with much higher initial intensity error than any other forecast model (not shown). In addition, the growth rate in mean absolute intensity error among the global forecast models is similar through the first 48 hr. Overall, the GFDN features the lowest mean absolute intensity error among the models considered here and is very similar to the forecast performance of the JTWC beyond 36 hr.

3.3.3. Regional Outlooks of Tropical Cyclone Activity

We now consider the VarEPS's ability to produce skillful regional outlooks of TC activity (i.e. the probability that a tropical depression strength or greater vortex will be located within a region) by dividing the NIO into two sub-domains: the Arabian Sea and the Bay of Bengal. Figure 3.7a shows the ROC for the Arabian Sea using the VarEPS forecasts from April–June and August–December of 2007–2010. Similar to the ROC analysis of the VarEPS TC genesis forecasts, the VarEPS forecasts for TC activity in the Arabian Sea exhibit moderate probabilities of detection with very low false alarm rates, even as the forecast decision threshold decreases to lower probabilities. As forecast lead-time increases, a transition in forecast performance occurs after a lead-time of ten days. The ROC curve for forecasts with lead-times five to ten days in advance have similar POD and FAR with forecasts at shorter lead-times. However, comparing the ROC curve for forecasts with lead-times of ten to fifteen days relative to five to ten days in advance

reveals a large decrease in the probability of detection although the false alarm rate essentially remains constant. From this analysis it is unclear whether this decrease in POD is due an inherent lack of predictability at this longer-time scale or is functionally dependent on the current configuration of the VarEPS with reduced horizontal resolution at lead-times greater than 240 hr. It is clear that the current configuration of the VarEPS is incapable of generating forecast probabilities of TC activity in the Arabian Sea greater than 30% for lead-times of ten to fifteen days—one reason why the POD is substantially lower than for forecasts with lead-times less than ten days.

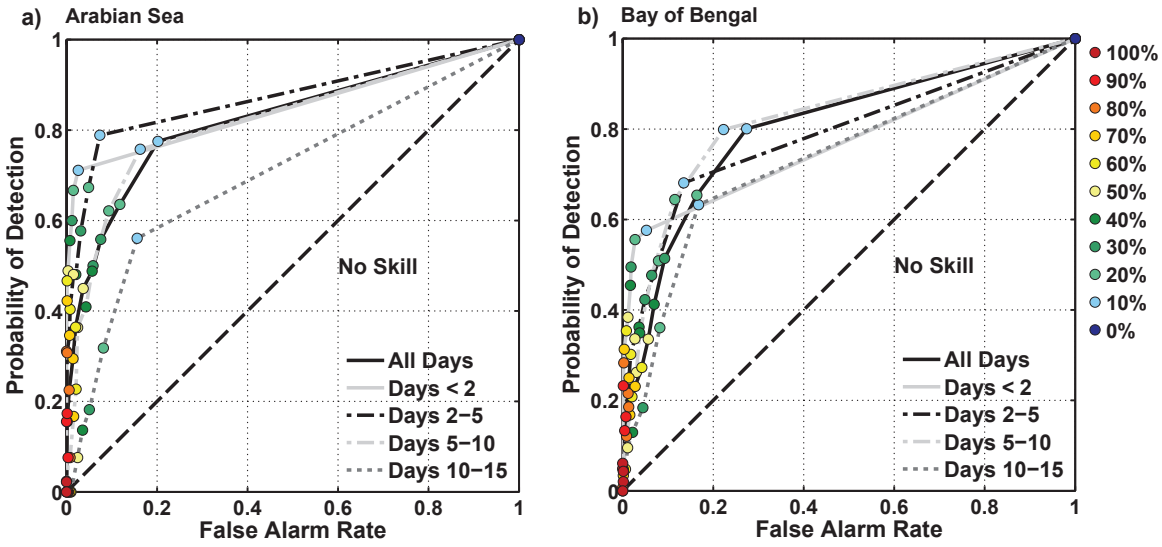

Figure 3.7: Relative operating characteristic for the a) Arabian Sea and b) Bay of Bengal using the VarEPS forecasts from 2007 to 2010 during the months of April–June and August–December for various forecast lead-times.

Figure 3.7b shows the ROC for the Bay of Bengal using the same set of VarEPS forecasts as in Figure 3.7a. Relative to the Arabian Sea, the VarEPS's forecasts of TC activity in the Bay of Bengal have lower probabilities of detection for shorter lead-times, but higher POD values at longer lead-times. The relative change in POD and FAR with forecast lead-time indicates that overall the VarEPS forecasts are more skillful in the Bay of Bengal than the Arabian Sea at extended lead-times.

To quantify how well the VarEPS forecasts of TC activity perform in the Arabian Sea and the Bay of Bengal as a function of forecast lead-time, two skill score metrics are used: the Brier skill score (BSS) and the ROC score. The BSS measures the accuracy or relative skill of a forecast over climatology by comparing whether or not an event is forecast to occur relative to observations. A BSS greater than zero implies forecast skill beyond climatology. In this analysis, a 30-yr climatology (1980–2009) of tropical cyclones was developed from the JTWC best-track dataset. Although the BSS metric reveals how skillful a forecast system is relative to climatology, the BSS is regarded as a harsh forecast standard as it can often hide useful formation information even when the BSS is less than zero (Mason 2004). Therefore, we use the ROC score as another skill metric. Table 3.1 provides the BSS and ROC scores for the Arabian Sea and the Bay of Bengal based on the VarEPS forecasts from April–June and August–December from 2007 to 2010 as a function of forecast lead-time. It should be noted that a BSS/ROC score of one indicates a perfect set of forecasts. To establish statistical significance at the 95% confidence level, a nonparametric bootstrap test was used. In this case, statistical significance of the BSS (ROC score) is determined if the 95% confidence interval of the BSS exceeds zero (0.50). From Table 3.1, the BSS metric indicates that the VarEPS forecasts for TC activity in the Arabian Sea are skillful beyond climatology for forecasts up to ten days in advance. Beyond ten days, however, forecasts for TC activity are not skillful relative to climatology. In contrast to the BSS metric, the ROC score is well-above 0.50 out to 15 days, which indicates that the VarEPS provides skillful forecasts of TC activity through two weeks in the Arabian Sea relative to random-chance alone.

Table 3.1 also provides the BSS and ROC scores for TC forecasts in the Bay of Bengal. According to the BSS, forecasts less than five days in advance in the Bay of Bengal are nearly 60% less skillful than in the Arabian Sea, which is likely due to the higher frequency of false alarms in the Bay of Bengal (c.f. Figure 3.4c and Figure 3.7a-b). However, consistent with the ROC interpretation of Figure 3.7a-b, the BSS is positive

and statistically greater than zero at the 95% confidence level through a lead-time of ten days. For the ten to fifteen day forecast period, the BSS indicates that the VarEPS performs as well as climatology in the Bay of Bengal. Although TC forecasts from the VarEPS in the Arabian Sea are more skillful than those in the Bay of Bengal for lead-times less than ten days, this result does not hold at longer time scales where the Bay of Bengal forecasts have on average an 18% higher BSS. If the ROC score metric is used to determine forecast skill, the VarEPS forecasts for TCs in the Bay of Bengal are skillful through fifteen days, which is similar to the ROC score results for the Arabian Sea.

Table 3.1: Brier skill scores (BSS) and relative operating characteristic scores for the Arabian Sea and the Bay of Bengal based on VarEPS forecasts for tropical cyclone activity during the months of April–June and August–December for 2007–2010. BSS (ROC scores) in bold are statistically different from 0 (0.5) at the 95% confidence level.

Arabian Sea	Brier Skill Score	Relative Operating Characteristic Score
All Forecast Days	**0.17**	**0.82**
Forecast Days ≤ 2	**0.47**	**0.85**
Forecast Days 2–5	**0.32**	**0.87**
Forecast Days 5–10	0.04	**0.82**
Forecast Days 10–15	-0.14	**0.70**
Bay of Bengal	Brier Skill Score	Relative Operating Characteristic Score
All Forecast Days	**0.09**	**0.80**
Forecast Days ≤ 2	**0.30**	**0.77**
Forecast Days 2–5	**0.16**	**0.79**
Forecast Days 5–10	**0.16**	**0.82**
Forecast Days 10–15	-0.02	**0.74**

The results in Table 3.1 may be compared to a similar analysis that Belanger et al. (2010) performed for the tropical North Atlantic using the ECMFS during the hurricane seasons of 2008 and 2009. This analysis found that that the most predictable region for TC activity in the North Atlantic is the MDR, as ROC scores for forecast days 8–14 are 0.81 and 0.75 for forecast days 15–21. Relative to these findings, the results presented here indicate similar predictability in the Bay of Bengal as in the NATL MDR, with TC predictability in the Arabian Sea more characteristic of that of the Caribbean Sea at extended forecast lead-times. Belanger et al. (2010) attribute the predictability at these time scales to the ability of the VarEPS to forecast accurately the magnitude of deep-layer (850-200 hPa) vertical wind shear as well as the correct frequency of pre-existing cyclonic vortices such as AEWs. However unlike the Caribbean Sea, where predictability is modulated largely by the variability in intensity and location of the tropical upper-tropospheric trough, in the Arabian Sea it is likely the combination of dry environmental air and changes in deep-layer vertical wind shear in association with the onset and end of the south Asian monsoon. In addition, Vitart (2009), along with Belanger et al. (2010), shows that regional TC predictability in the tropical Atlantic is strongly modulated by the phase and amplitude of the MJO. Given the more pronounced impact that the MJO has in the NIO (Webster and Hoyos 2004, Hoyos and Webster 2007, Krishnamohan et al. 2012), we expect this finding to be even more applicable to TCs in the NIO.

3.4. Summary and Conclusions

The performance of the ECMWF VarEPS in forecasting tropical cyclones in the North Indian Ocean has been examined for the period 2007–2010. The VarEPS is shown to have low false alarm rates and moderate to high probabilities of detection for forecast lead-times through seven days. The VarEPS TC genesis forecast performance is sensitive to the time window that is used to define whether or not an event is forecast to

occur. Based on an optimization procedure to achieve the highest ROC score (i.e. maximizing probability of detection and minimizing the false alarm rate), the optimum forecasting combination to predict TC genesis is a 48-hr time window for a forecast lead-time through five days. Thereafter, the full time period of the VarEPS integration should be used to generate TC genesis forecasts five to ten days in advance.

Analysis of the VarEPS forecasts from 2007 to 2010 shows that tropical cyclone track forecasts made prior to TC genesis perform 15–30% better than track forecasts produced after TC genesis has occurred. For a lead-time of 24 hr to 240 hr, the total mean track error grew at a rate 41 n mi per day such that by a lead-time of 120 hr (240 hr), the average track error of all VarEPS forecast is 224 n mi (409 n mi) with a 50% interval of 116–324 n mi (186–498 n mi). The performance of these track forecasts prior to TC genesis is remarkable considering that the IMD's average 72-hr forecast track error is 500 km (270 n mi), and these forecasts are issued only after TC genesis has occurred. In terms of post-genesis TC track forecasting, the VarEPS forecasts for a lead-time of 24 hr average 71 n mi with a 50% interval of 35 to 98 n mi, and by 120 hr the mean track error is 325 n mi with a 50% interval of 148–427. In addition, both the pre-genesis and post-genesis track analyses show that the distribution of ensemble track error becomes increasingly non-Gaussian as forecast lead-time increases. Therefore, to maximize forecast track predictability using the VarEPS, the full distribution of VarEPS track forecasts should be considered and not the mean VarEPS alone. Although the VarEPS control and ensemble mean forecast on average starts with the largest initial track error when compared to other forecast models (i.e. NOGAPS, UKMET, GFS, GFDN), for forecast lead-times of 12 hr and beyond, the VarEPS control and ensemble mean show the lowest track errors among all other model forecasts. In fact, the VarEPS control and ensemble mean on average exhibit slightly lower track errors than the JTWC through a lead-time of 72 hr. Since these forecasts are not bias-adjusted to account for the difference in the starting location of the observed TC relative to the initialized location in

the VarEPS model or systematic along-track or cross-track biases, additional statistical post-processing steps could be applied to significantly lower the average track errors of the VarEPS at extended lead-times (see Chapter 5).

In addition, the VarEPS track forecasts were translated into regional outlooks to provide the likelihood of TC activity in the Arabian Sea and the Bay of Bengal. Skill score metrics including the Brier skill score and the relative operating characteristic score were used to evaluate the VarEPS forecasts. The BSS statistic indicates that the VarEPS TC forecasts for the Arabian Sea and Bay of Bengal are predictable relative to climatology through ten days, whereas the ROC score statistic shows that TC activity for both regions are predictable through two weeks.

Based on this evaluation of the VarEPS TC forecasts, it appears feasible for warning agencies in the NIO to begin providing a probabilistic TC formation outlook that assesses the potential for TC development through a lead-time of seven days. When the probability of formation is within moderate (30–60%) levels, the VarEPS's probability of detection will average around 60% with a false alarm rate of about 30% for a lead-time of seven days. In addition, since the distribution of the VarEPS forecasts provides a dynamical measure of the forecast uncertainty in the atmosphere's future state (Dupont et al. 2011), some TCs will be more predictable than others. Therefore, operational forecasts could include a probabilistic outlook including both TC track and intensity information derived from the VarEPS. Although the VarEPS tends to be slightly underdispersive at longer forecast lead-times (Majumdar et al. 2010), some additional statistical post-processing steps including bias-correction and probability calibration could be incorporated to ensure that the final forecast track and intensity probabilities are well-conditioned relative to observations (see Chapter 5).

CHAPTER 4
TROPICAL CYCLONE PREDICTABILITY IN THE NORTH ATLANTIC OCEAN ON INTRASEASONAL TIME SCALES

The work presented in this chapter is published in Monthly Weather Review (Belanger et al. 2010).

4.1. Background and Motivation

As discussed in Section 1.2.3, intraseasonal TC forecasts have only recently been examined especially those forecasts generated from dynamically-based NWP systems. To explore how phasing of the MJO impacts tropical cyclone frequency in the ECMFS, Vitart (2009) conducted hindcast experiments from 1989 to 2008 using an ensemble of 15 members that were initialized on the 15th of each month and ran for 46 days. Using the Wheeler and Hendon (2004) MJO Index, the model simulations showed an increase (decrease) of TC activity across the Atlantic during MJO Phases 2–3 (6–7), which is in qualitative agreement with observations but with a weaker MJO impact than is seen. Vitart (2009) also showed that MJO phasing projects onto the risk of landfall, as determined by the accumulated cyclone energy over land, with MJO Phases 2–3 showing a higher risk of landfall than MJO Phases 6–7. Although Vitart's study highlights the ability of the dynamical-modeling system to produce intraseasonal TC variability based on MJO phasing in agreement with observations, additional work is required to understand the source of TC predictability on intraseasonal time scales and the current limitations of these extended forecasts.

In this chapter, we evaluate the forecast skill of intraseasonal TC predictions for the tropical North Atlantic using the ECMFS. Since the ECMFS can reproduce the observed variability of the MJO realistically through three weeks (Vitart 2009), we expect that this system is capable of reliably identifying periods in a forecast cycle that

should portend higher or lower TC activity than normal. After introducing the data and methodology in Section 4.2, results from the ECMFS during the 2008 and 2009 Atlantic hurricane seasons are presented along with findings of a predictability analysis of the large-scale environment in Section 4.3. We close with a summary of the potential implications from this research along with recommendations on how the ECMFS may be used to generate skillful intraseasonal TC forecasts in Section 4.4.

4.2. Data and Methodology

This study examines North Atlantic tropical cyclones and the large-scale environment within which they are embedded for the period June to October of 2008–2009. The analysis employs forecasts from the ECMFS, the ERA-Interim reanalysis, and observations of tropical cyclone tracks from HURDAT. A complete description of the ECMFS is found in Section 2.2.2. To extract tropical cyclones from the model forecast fields, the track algorithm of Vitart et al. (1997) is used (see Section 2.3.2 for more information on the tracking methodology). In addition to tracking TCs in the ECMFS, the trajectory of AEWs is analyzed using a Hovmöller method detailed in Agudelo et al. (2011). The easterly wave tracking algorithm uses 2 – 6 day Fourier filtered, westward-moving 850 hPa relative vorticity averaged for the latitudinal band of 5^{o}N–15^{o}N. Based on the timing and location of the local maximum of 850 hPa relative vorticity that pass westward across 20^{o}W, a recursive algorithm is applied to identify the longitudinal extent of the easterly wave as a function of time. The working hypothesis is that the forecast frequency of AEWs modulates the predicted levels of tropical cyclone activity in the ECMFS and is another source of intraseasonal TC predictability besides the MJO. Intraseasonal TC predictability is assessed using the correlation coefficient, Brier skill score, relative operating characteristic score, and the reliability diagram (see Section 2.4 for more information on these verification tools).

4.3. Results and Discussion

Tropical cyclone activity predicted from the ECMFS is compared with TC climatology (1970 – 2000) for the forecasts initialized on 7 August 2008 and 14 August 2008 (Figure 4.1a,b). For the 7 August 2008 forecast, lower than normal TC activity was projected to occur across much of the North Atlantic. However, the monthly forecast issued just one week later indicated that the subsequent 32-day forecast period would be quite active, especially across much of the northern MDR and the Greater Antilles. Observations from HURDAT for this time period indicate that five TCs developed (i.e. Tropical Storm Fay, Hurricane Gustav, Hurricane Hanna, Hurricane Ike, and Tropical Storm Josephine) and moved through the tropical North Atlantic region that was forecast on 14 August 2008 to see higher than normal TC activity but one week prior was forecast to be below normal.

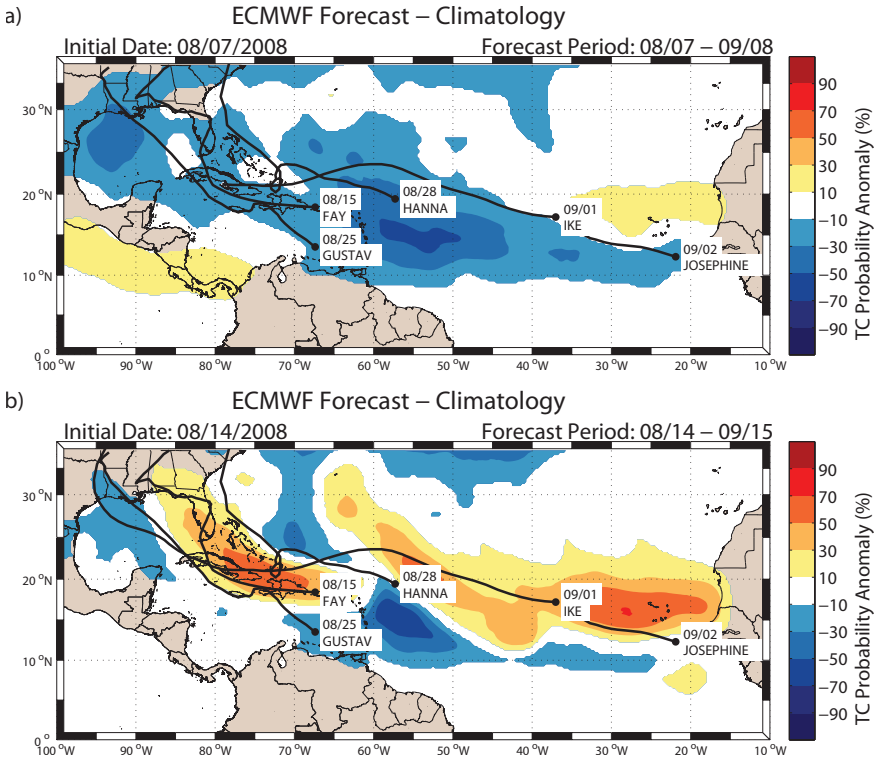

Figure 4.1: (a) Variance from climatology (1970–2000) of TC probabilities in the ECMFS for the full 32-day period of monthly forecasts initialized on 7 August 2008 and one week later (b) on 14 August 2008. The observed tracks of tropical cyclones that occurred during each 32-day period are overlaid in black.

To explain the weekly variability in predicted levels of TC activity in the ECMFS, we hypothesize that both the phasing (Vitart 2009) and the amplitude of the MJO modulates the projected levels of TC activity in the model and ultimately regulates which intraseasonal predictions will be more skillful than others. In the case of Figure 4.1a (7 August 2008), the MJO projection onto the Wheeler-Hendon phase space was weak and centered in the Indian Ocean. However, the convectively active phase on 14 August 2008 was strong and remained centered in the Indian Ocean, which provides anecdotal evidence of this forecast sensitivity. This hypothesis is evaluated more formally in Section 4.3.3.

4.3.1. Skill of the ECMWF Monthly Forecast System

To provide an evaluation of the ECMFS skill, spatial BSS are shown in Figure 4.2a-d for all of the monthly forecasts produced weekly from June to October during 2008 and 2009. For each composite, the reference forecast is the climatology of tropical cyclone activity from 1970–2000 using HURDAT and varies based on the specific time period covered by each monthly forecast. To show the temporal sensitivity of the forecast skill, the analysis for the monthly forecast period is divided by the number of weeks-in-advance (e.g., Week 1 includes forecast days 1–7; Week 2: days 8–14; Week 3: days 15–21; Week 4: days 22–28). Note that the ECMWF defines their weekly periods differently so that the full 32-day forecast is covered by a four-week period with the same model resolution (Vitart et al. 2008). From Figure 4.2a and Figure 4.2b, it is seen that the ECMFS produces skillful predictions of tropical cyclone activity across much of the tropical North Atlantic including the MDR, West Atlantic (20^{o}N–35^{o}N, 60^{o}W– 90^{o}W), northern Caribbean Sea, and the Gulf of Mexico. For Week 3, the regions of positive Brier skill scores narrows and is focused in the southern Gulf of Mexico and central MDR (Figure 4.2c). By Week 4, the central MDR is the only region of the tropical North Atlantic that features skill above climatology (Figure 4.2d). Finding skill especially for

Weeks 3 and 4 in the MDR is unexpected since the major source of predictability beyond two weeks is the MJO, and the observed impact of MJO variability on TC activity is focused in the Gulf of Mexico and western Caribbean Sea (Maloney and Hartmann 2000a,b, Klotzbach 2010). Possible explanations for this long-lead forecast skill in the central MDR are provided in Section 4.3.2.

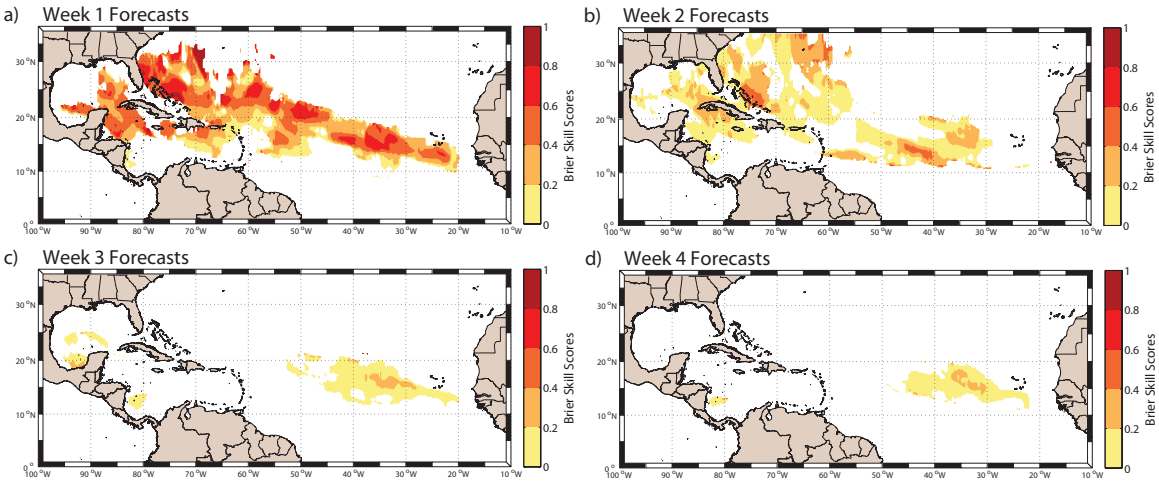

Figure 4.2: Brier skill scores for the ECMFS as a function of weeks-in-advance. The weekly composites include all of the monthly forecasts made from June to October during 2008 and 2009. Weeks 1, 2, 3, and 4 cover days, 1–7, 8–14, 15–21, 22–28, respectively. The tropical cyclone climatology, which is determined using HURDAT for the period 1970–2000, is the reference forecast. Values above zero indicate skill beyond climatology, and a value of one implies a perfect forecast. Note the Brier skill scores have been smoothed once by a 9-point running mean.

The conditional bias of the probabilistic forecasts is investigated using the reliability diagram along with the occurrence frequency of each forecast probability level (Figure 4.3). The ECMFS is overconfident (i.e., TC forecast probabilities too high) for all probability levels and across all weekly periods. Although the composite reliability of the Week 1 forecasts increases linearly as the forecast probability level increases, the composite reliability of the Week 2–4 forecasts reaches a plateau near 40–50% and only agrees with observations approximately 20% of the time. Since the occurrence frequency of forecast probability levels rarely exceeds 50% for these longer lead-time forecasts (i.e., Weeks 2–4), it is unclear whether this reliability plateau represents a dynamical

predictability limit or is due simply to an insufficient sample size of monthly forecasts. Unlike a climatology forecast that provides no resolution (Wilks 1995), the ECMFS provides moderate resolution in that it can discriminate between TC events and nonevents across all weekly periods.

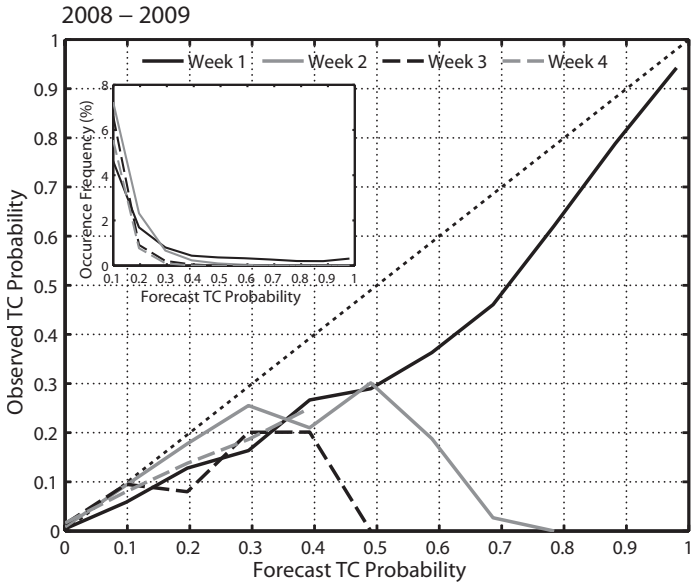

Figure 4.3: Reliability diagram of the ECMFS as a function of weeks-in-advance. The inset has the relative occurrence frequency of each forecast TC probability level normalized by the size of the 0.25^{o} x 0.25^{o} grid domain. The reliability diagram is constructed using the full domain of the probability forecasts along with observations.

The ECMFS's ROC for TC activity complements the reliability diagram. Given the spatial and temporal differences in Brier skill scores (Figure 4.2a-d), the ROC is calculated regionally using the maximum TC forecast probability as a function of weeks-in-advance in four regions: the Gulf of Mexico, Caribbean Sea, West Atlantic, and Main Development Region (Figure 4.4a-d). Table 4.1 provides the regional ROC scores for Weeks 1–4. The ROC scores are calculated by integrating the hit rates for each region as a function of increasing false alarm rates. A ROC score of 0.5 indicates no forecast skill and a value of 1.0 characterizes a perfect prediction. For the Week 1 forecasts, the Gulf of Mexico, West Atlantic, and Main Development Region have the highest ROC scores (0.85–0.87) with low false alarm rates and high probabilities of detection as a function of

increasing probability level (Figure 4.4). For the Week 2 forecasts, the West Atlantic and Main Development Region maintain moderate to high forecast skill with ROC scores around 0.8, while the forecasts for TC activity in the Gulf of Mexico have the largest decrease from Week 1 to Week 2 (from 0.85 to 0.66). For the Week 3 and 4 forecasts, ROC scores average around 0.65 for the Gulf of Mexico, Caribbean Sea, and West Atlantic with higher forecast skill in the Main Development Region (ROC score ~ 0.75). Since the ECMFS ROC scores remain above 0.5 through four weeks, this result indicates that forecast skill exists across all regions and for all time periods considered even though the BSS analyses in Figure 4.2 imply that the MDR and Gulf of Mexico are the only regions that feature forecast skill on time scales longer than two weeks.

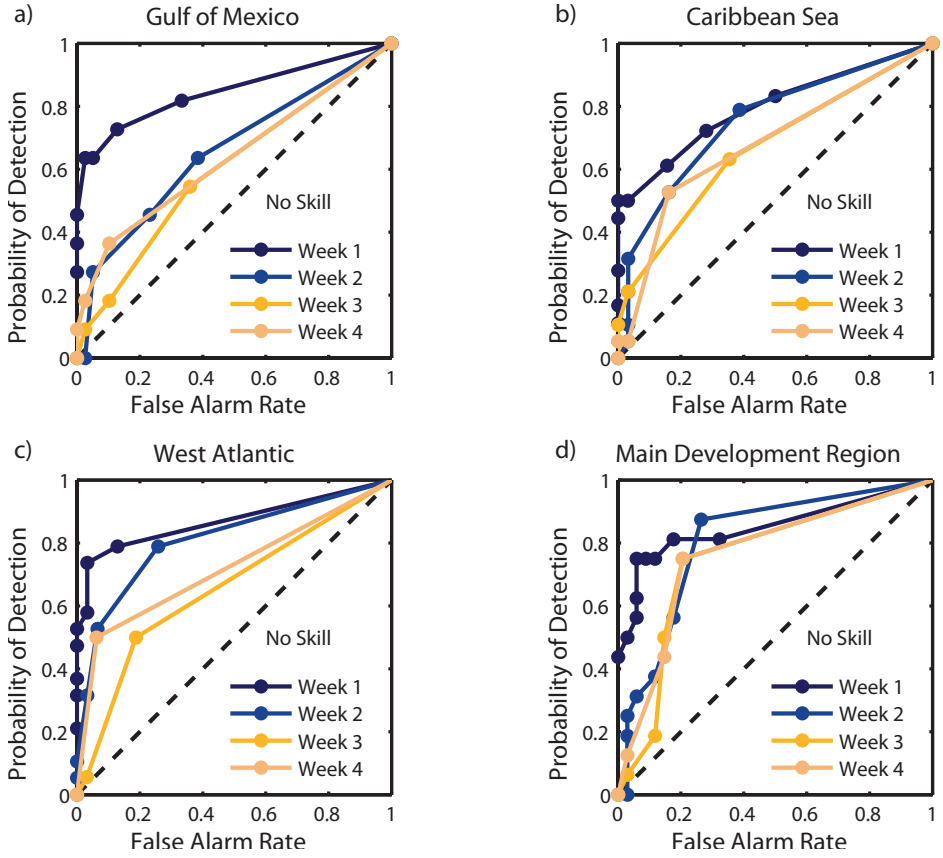

Figure 4.4: Relative operating characteristic of the ECMFS for the maximum TC forecast probabilities in the (a) Gulf of Mexico, (b) Caribbean Sea, (c) West Atlantic (20°N–35°N, 60°W–90°W), and (d) Main Development Region (10°N–20°N, 20°W– 60°W) as a function of weeks-in-advance. The diagonal, dashed line indicates no forecast skill.

Table 4.1: Relative operating characteristic (ROC) scores for TC probability forecasts within various regions of the North Atlantic including the Gulf of Mexico, Caribbean Sea, West Atlantic, and Main Development Region as a function of weeks-in-advance. Higher ROC scores indicate greater forecast skill, where 0.5 is the threshold for no forecast utility.

Relative Operating Characteristic Score	Week 1	Week 2	Week 3	Week 4
Gulf of Mexico	0.85	0.66	0.60	0.64
Caribbean Sea	0.80	0.75	0.67	0.68
West Atlantic	0.87	0.81	0.65	0.72
Main Development Region	0.85	0.81	0.75	0.76

4.3.2. Predictability of the Large-Scale Environment

Since large-scale environmental variables explain at least 50% of the total variance in TC genesis events across the tropical North Atlantic (Agudelo et al. 2011), the predictability of the large-scale environmental flow is evaluated to explain why forecast skill utilizing the ECMFS exists on intraseasonal time scales across the tropical Atlantic. The analysis focuses on the ability of the model to capture the variability in deep-layer vertical wind shear (850–200 hPa), the frequency of AEWs, and the variance in 850 hPa relative vorticity. Other environmental variables such as column integrated humidity, upper (lower) level divergence (convergence), large-scale vertical velocity, or mid-latitude trough interactions may also be relevant, but their importance remains a topic of future work.

Spatial correlation coefficients between forecast values of deep-layer vertical wind shear and observations as determined by the ERA-Interim reanalysis (see Section 2.1.1 for more information) are given in Figure 4.5a-d as a function of weeks-in-advance. The correlation coefficients are calculated using the mean 12-hr correlation coefficients during each week of the monthly forecast period. Correlation coefficients exceed 0.8 for

most regions of the tropical Atlantic including the MDR, West Atlantic, and the Gulf of Mexico for Week 1. In Weeks 2–4, the Gulf of Mexico and the MDR have the highest correlations, as the forecast system captures one-third of the total variance in vertical wind shear. On the other hand, the lowest predictability is found in the eastern Caribbean Sea. Given the close correspondence between regions of positive BSS, high ROC scores, and moderate to high correlation coefficients for vertical wind shear, this result suggests that forecast skill for tropical cyclones especially in the central MDR and West Atlantic may be due to accurate vertical wind shear forecasts.

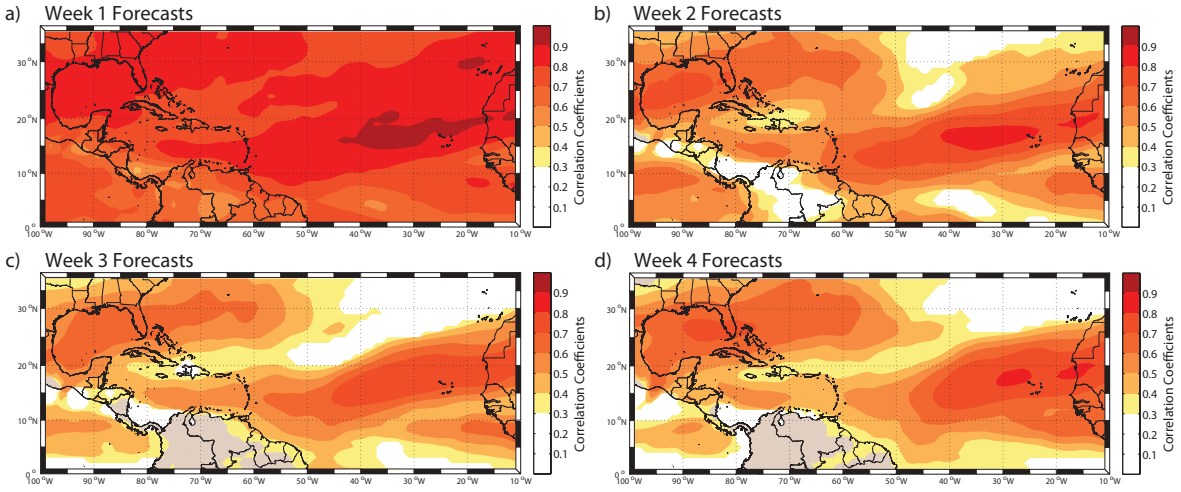

Figure 4.5: Correlation coefficients between the ECMFS and the ERA-Interim Reanalysis for deep-layer (850–200 hPa) vertical wind shear as a function of weeks-in-advance. The weekly composites include all of the monthly forecasts made from June to October of 2008 and 2009. Weeks 1, 2, 3, and 4 cover the days, 1-7, 8-14, 15-21, 22-28, respectively. Shaded regions are statistically significant at the 95% confidence level using a bootstrap resampling method.

The inability of the ECMFS to reproduce the observed variance in vertical wind shear in the Caribbean Sea, as seen in Week 3 and Week 4 (Figure 4.5c-d), is tied to variability in the intensity of the tropical upper-tropospheric trough, as this region exhibits more variance in vertical wind shear than any other location in the tropical North Atlantic (not shown). Vertical wind shear is thought to inhibit TC genesis and intensification by increasing ventilation (Gray 1968), by modifying vertical stability

(DeMaria 1996), and by inducing changes in the secondary circulation (Bender 1997). The relationship between vertical wind shear variations in the tropical Atlantic and TC frequency on intraseasonal time scales has been discussed by Mo (2000), Maloney and Shaman (2008), and Vitart (2009).

The impact of AEW frequency on tropical North Atlantic TC predictions by the ECMFS is evaluated as a potential contributor to the intraseasonal forecast skill. The application of the modified AEW tracking scheme by Agudelo et al. (2011) to the ECMFS from 2008 and 2009 shows that the impact of AEW variability is largest for most of the MDR and throughout the Gulf of Mexico (Figure 4.6a). For these regions, about 25% of the total variance in tropical cyclone activity may be attributed to the forecast frequency of AEWs. In addition, the spatial pattern of covariability coincides with the regional locations that have skillful TC forecasts (i.e., positive BSS and high ROC scores), which suggests that the forecast frequency of AEWs is an important component in the intraseasonal predictions of TCs in the ECMFS.

The ability to reproduce the variability in low-level relative vorticity is inherent to the predictability of the large-scale environmental flow, as ambient vorticity is a necessary condition for TC seedling growth and development especially from AEWs (Berry and Thorncroft 2005). The correlation coefficients between the predicted variance in 850 hPa relative vorticity in the MDR and the observed variance from the ERA-Interim reanalysis are shown in Figure 4.6b-d for the years 2008–2009. During June–October 2008, the correlation coefficients drop below statistically significant values (< 95% confidence level) after day 10, but then increase from zero to about 0.45 during Weeks 3–4. Whereas the predictability of the variance in 850 hPa relative vorticity for the MDR was limited to Week 1 during 2008, variability in low-level relative vorticity for 2009 was much more predictable (out to four weeks) especially across much of the eastern MDR, which coincides with a region that had positive BSS (Figure 4.2) and moderate to high ROC scores (Figure 4.4d). This interannual difference may reflect the

impact of the El-Niño Southern Oscillation (ENSO) on tropical Atlantic predictability, as El Niño conditions prevailed throughout much of the 2009 hurricane season, whereas ENSO neutral conditions occurred during 2008.

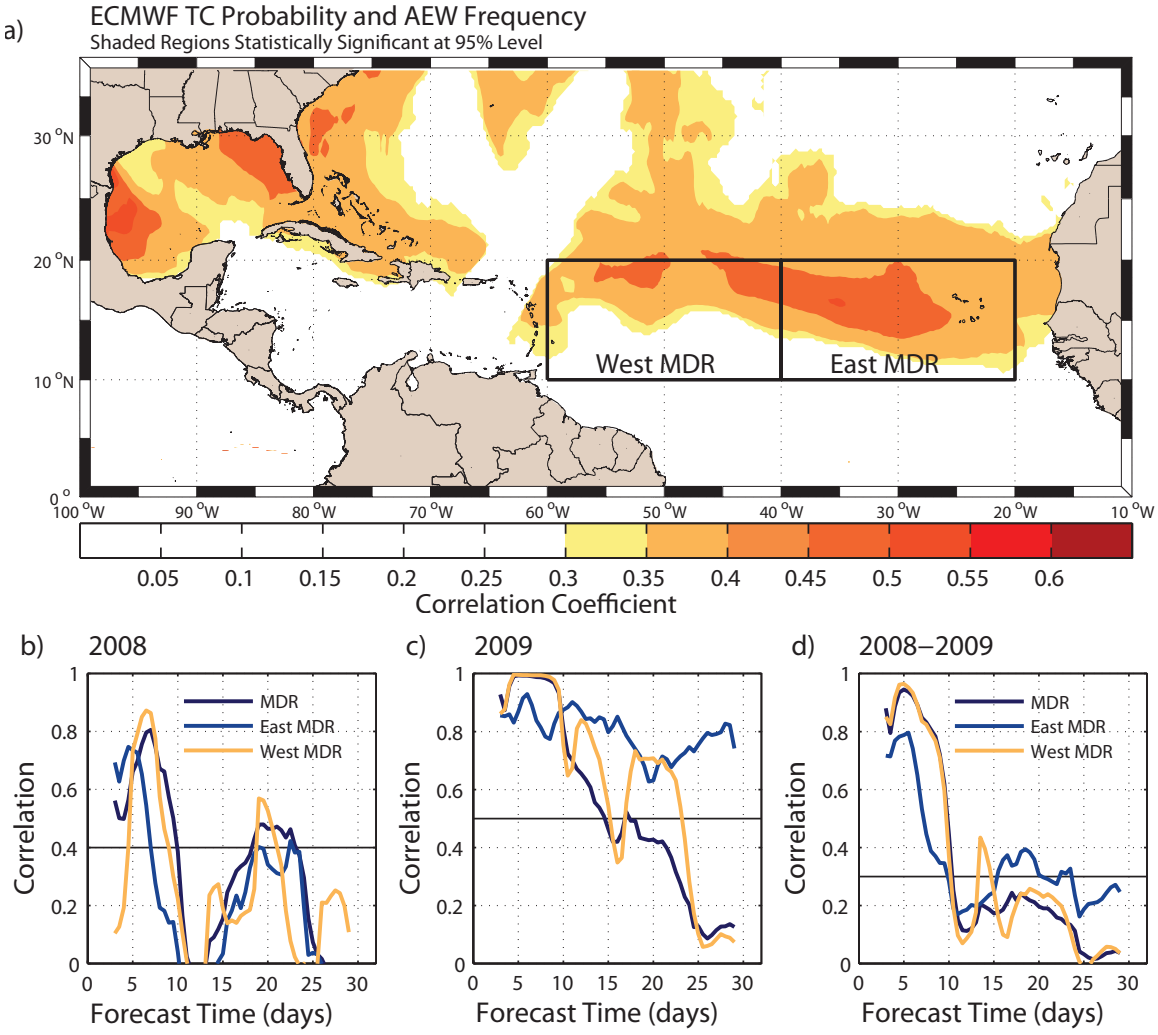

Figure 4.6: (a) Spatial correlation coefficients between forecast probability levels of tropical cyclone activity and the frequency of easterly waves in the ECMFS. Shaded regions are statistically significant at the 95% confidence level using a bootstrap resampling method. Time series of the correlation coefficients between the ECMFS and the ERA-Interim reanalysis for 850 hPa relative vorticity variance as a function of days in advance in the MDR, East MDR, and West MDR for (b) 2008, (c) 2009, and (d) 2008–2009.

4.3.3. Relationship with the Madden-Julian Oscillation

Forecast sensitivity to the phasing and intensity of the MJO is examined as another source of variability in the TC forecasts from the ECMFS. Mean TC probability forecasts of the ECMFS based on the location at the time of model initialization of the MJO in the Indian Ocean (Phases 1–3), West Pacific (Phases 4–5), and the East Pacific/Atlantic (Phases 6–8) are shown in Figure 4.7a-c, respectively, provided the MJO had an amplitude of at least one standard deviation. Since the impact of MJO phasing in the model forecast is sensitive to the seasonal climatology of TC activity across the North Atlantic, this impact is accounted for in Figure 4.7d–f, which are the mean TC forecasts after the 32-day TC climatology for each monthly prediction has been removed. When the MJO is centered in the Indian Ocean (Figure 4.7a) at the time of model initialization, the ECMFS will typically predict elevated levels of TC activity during the following 32-day period in a region emanating from the MDR and stretching into the Caribbean Sea, West Atlantic, and Gulf of Mexico. If these cases are adjusted by the seasonal climatology of TCs (e.g., Figure 4.7d), the anomalous elevation in TC activity is confined mainly to the MDR and north of the Greater Antilles. For time periods when the MJO is located in the West Pacific (Figure 4.7b), elevated levels of TC activity across the tropical Atlantic are reduced and confined to the western Caribbean and the eastern MDR. When adjusted by the seasonal TC climatology, the predicted level of TC activity across much of the West Atlantic is suppressed relative to climatology (Figure 4.7e). When the MJO is initially located in the East Pacific/Atlantic (Figure 4.7c), elevated TC activity is confined only to the western Caribbean Sea and the southern Gulf of Mexico, which is consistent if the TC forecasts are adjusted by the seasonal climatology (Figure 4.7f). These forecast tendencies are in agreement with previous observational and modeling studies on the relationship between MJO phasing and Atlantic TC activity (Mo 2000, Barrett and Leslie 2009, Vitart 2009, Klotzbach 2010).

If the sensitivity analysis to MJO phasing is completed for the ECMFS forecasts when the MJO amplitude is less than one standard deviation, no significant departure from TC climatology is found across the tropical Atlantic except for MJO Phases 4–5 in which TC activity in the MDR, Caribbean Sea, and Gulf of Mexico is well below climatology (not shown). These results indicate that in addition to MJO phasing (Vitart 2009), the initial amplitude of MJO at the time of model initialization also modulates which regions of the tropical Atlantic will have enhanced or suppressed TC activity.

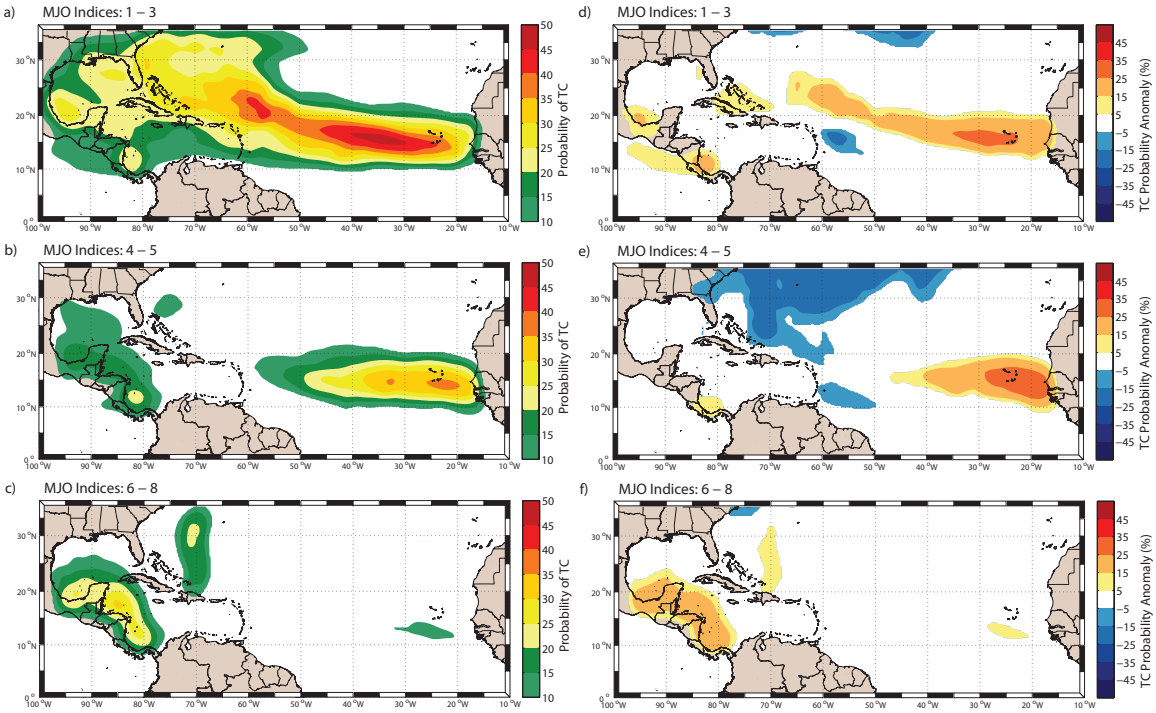

Figure 4.7: Mean TC probability forecasts for the entire 32-day period from the ECMFS initialized when the convectively-active phase of the MJO with amplitude of at least one standard deviation was centered in MJO (a) Phases 1 – 3 (Indian Ocean), (b) Phases 4 – 5 (Western Pacific), (c) Phases 6 – 8 (East Pacific/Western Hemisphere) at the time of model initialization. Mean TC probability forecasts as in panels a–c, except that the 32-day climatology of observed TC activity has been removed from each monthly forecast, for (d) Phases 1–3, (e) Phases 4–5, and (f) Phases 6–8. Note: The spatial probabilities have been smoothed once by a 9-point running mean.

From an operational perspective, knowing how the predicted levels of TC activity in the ECMFS will vary based on phasing of the MJO is insightful. However, this knowledge does not provide any guidance as to whether a forecast is reliable. Forecast

utility based on MJO phasing and intensity is displayed in Figure 4.8 in terms of the conditional reliability of the intraseasonal forecasts, which is based on regional clusters of the MJO location at the time of model initialization and for events when the amplitude of the MJO as determined by Wheeler-Hendon MJO index is at least one standard deviation. The most reliable forecasts (i.e., close proximity to the diagonal line) occur when the MJO is initially in the Indian Ocean (Phases 1–3). During the June to October period for 2008–2009, 42% of all the weekly forecasts were initialized when the MJO was initially centered in this region, and for this condition, the model provides excellent reliability for forecast TC probability levels from 0% to 40%. Between the forecast thresholds of 50% to 80%, the forecasts of Atlantic TC activity have no significant improvement over the 40% threshold. However, the probability levels from 80% to 100% tend to be more reliable, as they verify 50% to 80% of the time, respectively.

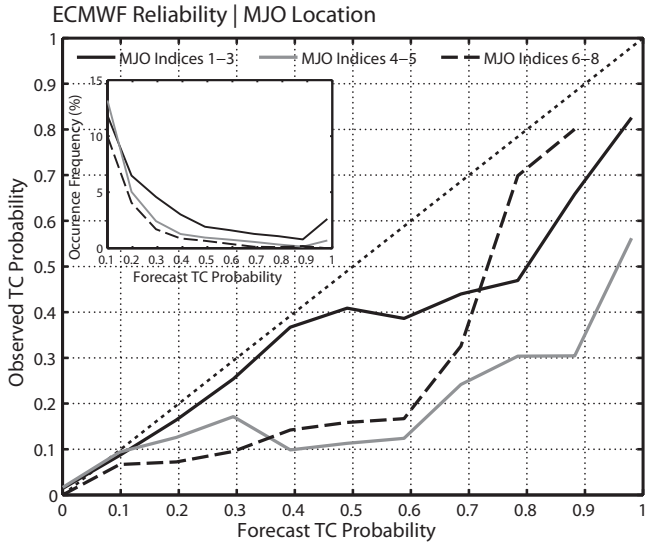

Figure 4.8: Conditional reliability of the TC probability forecasts from the ECMFS as a function of MJO location, at the time of model initialization. The inset is the relative occurrence frequency of each forecast TC probability level normalized by the size of the 0.25^{o} x 0.25^{o} grid domain. The reliability diagram is constructed using the full domain of the probability forecasts along with observed TC activity for each 32-day period.

Outside of the time periods when the MJO is initially located in the Indian Ocean, the reliability of the ECMFS forecasts of TC activity are significantly limited (Figure

4.8). When the MJO is either located in the western North Pacific (Phases 4–5) or the Western Hemisphere (Phases 6–8), forecast probability levels of TC activity between 0% and 60% are in agreement with observations less than 20% of the time, which in fact is even less reliable than a TC climatology forecast (not shown). Although reliability beyond the 60% forecast probability level increases, this apparent improvement in forecast utility is likely an artifact of the limited sample size for cases that exceed 60% (see Figure 4.8 inset). From this reliability analysis in conjunction with the spatial composites of TC activity conditioned on the phase of the MJO, Atlantic TC predictability on intraseasonal time scales using the ECMFS appears to reside during time periods when the MJO is convectively active and initially located in the Indian Ocean.

One caveat to these results is that they are based on only two years of data with one year (2009) having lower than normal MJO variability. The amplitude of the MJO during June–October 2008 was on average larger than in 2009 with the active convective phase located mostly in Phases 1–3 during 2008 (Figure 4.9a-b). Although these two seasons provide some indication on how Atlantic TC activity in the ECMFS may respond to the MJO, a larger sample of monthly forecasts that includes greater MJO variability along with unique background states (i.e., ENSO) is needed to confirm this sensitivity.

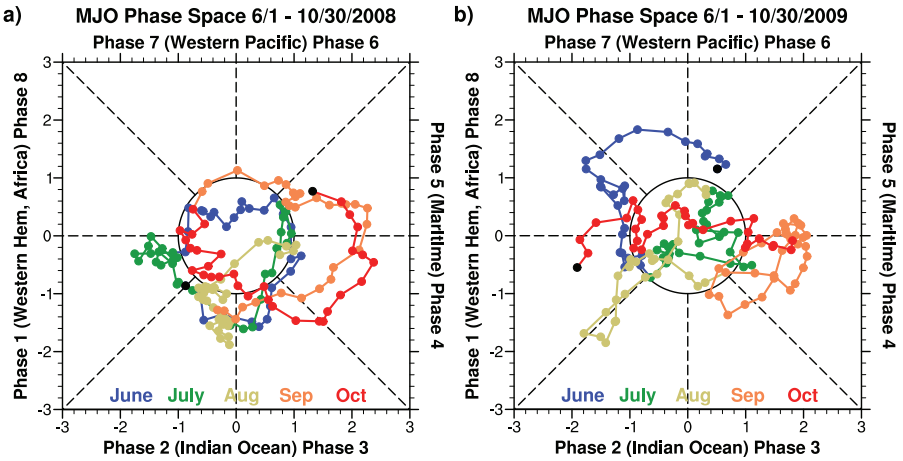

Figure 4.9: Phase space diagrams of the real-time multivariate MJO series 1 & 2 for the period 1 June through 30 October for (a) 2008 and (b) 2009. MJO data obtained from Australia's Bureau of Meteorology.

4.4. Conclusions

The predictability of TC activity in the tropical North Atlantic using the ECMWF Monthly Forecast System has been evaluated in terms of Brier skill scores, regional relative operating characteristic scores, and reliability diagrams to determine the conditional bias of the forecast scheme. From the BSS analysis, the tropical Atlantic regions that have forecast skill include the Main Development Region, West Atlantic, northern Caribbean Sea, and Gulf of Mexico during the first two weeks. For Week 3 forecasts, forecast skill is only found for the southern Gulf of Mexico and MDR, and at longer lead times the only skill is in the central MDR. Regional weekly ROC scores indicate that Week 1 forecasts have low false alarm rates for most regions of the tropical Atlantic (excluding the Caribbean Sea) and increasing probabilities of detection as the decision threshold is increased. Even at longer time scales, the four regions have forecast skill through Week 3 and Week 4 with average ROC scores of 0.75 for the MDR and 0.65 for the Gulf of Mexico, Caribbean Sea, and West Atlantic.

To evaluate why the ECMFS is able to produce skillful TC forecasts on intraseasonal time scales, the model's ability to forecast vertical wind shear is examined. For the Gulf of Mexico and MDR, the monthly forecasts explain nearly one-third of the observed variability in deep-layer vertical wind shear through four weeks. The regions of positive BSS and high ROC scores as a function of time coincide with the same regions as high vertical wind shear predictability.

We also determine whether the frequency of AEWs modulates TC predictions from the ECMFS. Although on interannual time scales AEW frequency is thought to only weakly impact North Atlantic TC frequency (Frank 1975, Thorncroft and Hodges 2001, Hopsch et al. 2007), around 25% of the total variance in TC activity predicted by the ECMFS is associated with the frequency of AEWs in the ECMFS. Furthermore, the spatial pattern of higher covariability values coincides with regional locations that feature skillful TC forecasts, including the Gulf of Mexico and the MDR, which indicates that

AEW frequency is an important contributor on intraseasonal time scales. Although this analysis has shown that vertical wind shear variability and easterly wave activity impacts Atlantic TC predictability in the ECMFS, additional research is still necessary to determine the role of other environmental variables such as column integrated diabatic heating, mid-level specific humidity, upper (lower) level divergence (convergence), and mid-latitude trough interactions.

Following Vitart (2009), phasing and amplitude of the MJO are considered to modulate TC predictions in the tropical Atlantic and, ultimately, their reliability. When the MJO is centered in the Indian Ocean at the time of model initialization, the ECMFS will typically predict elevated levels of TC activity during the 32-day forecast period in a region from the MDR into the West Atlantic. When the MJO is located in the West Pacific, the ECMFS predicts TC activity will be below climatology in the western North Atlantic. When the active convective phase of the MJO is in the Western Hemisphere, elevated TC probability levels according to the ECMFS are confined to the western Caribbean and southern Gulf of Mexico. These results are qualitatively consistent with Vitart (2009) who found MJO Phases 2–3 (6–7) at the time of model initialization cause an increase (decrease) of TC activity across the North Atlantic. When the reliability analysis of the ECMFS is conditioned on MJO location and amplitude, the most reliable TC forecasts from the model occur when the convectively active MJO is initially located in the Indian Ocean. The dynamical interpretation is that because the MJO propagates slowly, when the convectively active phase is over the Indian Ocean (Phases 2 and 3), favorable environmental conditions for TC formation will generally occur in the Atlantic for at least the first two to three weeks of the prediction. When the MJO is not centered in the Indian Ocean at the time of model initialization, reliability is significantly degraded as forecast probability levels between 0% and 60% verify less than 20% of the time with observations. Since the sensitivity analysis of the ECMFS to the MJO uses only two years of data with 2009 having lower MJO variability than 2008, additional research is

necessary to verify these results with a larger sample of monthly forecasts forced from unique climate states and with greater MJO variability.

The findings from the spatial skill analysis in conjunction with work by Vitart (2009) provide evidence that dynamically–based TC forecasts on intraseasonal time scales can be produced for the tropical North Atlantic. Further research is still needed to identify whether it is possible to extract additional TC predictability information from the MJO by improving how the ECMFS propagates a weaker MJO signal. Even considering the current dynamical model limitations, a statistical-dynamical method might be developed in which intraseasonal forecasts of TC activity in the tropical Atlantic would utilize these dynamical predictions during times when the MJO is convectively active and initially located in the Indian Ocean (see Chapter 5). During other time periods when the ECMFS is expected to be less reliable, statistical methods (e.g., Leroy and Wheeler 2008) could be applied. Since Maloney and Shaman (2008) have linked intraseasonal precipitation variability of the West African Monsoon and Atlantic ITCZ to TC frequency in the tropical Atlantic, further research is warranted to understand whether the ECMFS is capable of predicting this additional source of intraseasonal variability and its impact on tropical Atlantic TC activity.

CHAPTER 5
A SUBSEASONAL TROPICAL CYCLONE FORECAST SYSTEM FOR THE NORTH ATLANTIC OCEAN

5.1. Background and Motivation

Public and private sector entities are seeking extended-range forecasts of tropical cyclone formation, movement, maximum intensity, wind field structure, and associated impacts. Specifically, the energy industry has begun to exploit the extended predictability that exists for tropical cyclones in the North Atlantic to support decision making for energy marketing, sales and trading. Other industries such as retail and agriculture can make use of extended-range forecast information for logistical planning of goods and services based on the potential impacts of a landfalling TC. Given the increasing applications for extended-range tropical cyclone forecasts, the predictability studies discussed in Chapter 3 and Chapter 4 have been utilized to develop the essential components necessary for an extended-range tropical cyclone prediction system. The adopted forecast philosophy uses a hybrid approach in which dynamically-based ensemble forecasts from global numerical weather prediction centers are statistically-rendered to remove systematic model biases and to calibrate forecast spread. These ensembles are then used to generate deterministic and probabilistic forecasts for tropical cyclone track movement and maximum intensity change. The following chapter details the methodology that has been developed thus far and provides an evaluation of the operational forecasts using this approach for the North Atlantic hurricane seasons of 2009–2011.

5.2. Data and Methodology

The operational, twice daily outlook for North Atlantic tropical cyclones utilizes the ECMWF deterministic forecast at a horizontal resolution of 0.125^{o} for a lead-time through 240 hr as well as the 51-member ECMWF Variable Resolution Ensemble Prediction System at a horizontal resolution of 0.25^{o} through a lead-time of 240 hr and at a horizontal resolution of 0.50^{o} for a lead-time from 246 hr to 360 hr. After processing the ensemble members for tropical cyclones that are present in the initial conditions or that develop during the model integration, tropical cyclone tracks for each ensemble then undergo track and intensity calibration, which is discussed in Section 5.2.1 and Section 5.2.2. Ensemble interpretation using a high-predictable clustering approach is discussed in Section 5.2.3. Finally, in Section 5.2.4, the methodology used to calibrate and generate twice weekly intraseasonal forecasts for the North Atlantic is reviewed.

5.2.1. Tropical Cyclone Track and Probability Calibration

The first component of the ensemble forecast track correction process is the removal of initial position errors between the model initialization and observed location of an active tropical cyclone as determined by real-time invest coordinates provided by the National Hurricane Center through their ATCF system. Initial position errors are calculated in terms of latitudinal and longitudinal deviations between the model and observed tropical cyclone location and are subsequently removed for the initial time step and at subsequent lead-times for each ensemble member.

Each ensemble forecast track is then adjusted using cross-track and along-track errors from the historical forecast track error distribution derived from the ECMWF hindcasts. Here, the assumption has been made that model track biases in the ECMWF hindcasts, as a function of forecast lead-time, are representative of the real-time VarEPS. This assumption turns out to be valid for a variety of large-scale phenomena especially at extended forecast lead-times since the atmospheric model and ensemble generation

technique in the ECMWF hindcasts do utilize the real-time configuration of the VarEPS. However, since initial conditions for the hindcasts utilize the ERA-Interim reanalysis, which represents the frozen configuration of the ECMWF atmospheric model from late 2006 to early 2007, the hindcasts may not be completely representative of biases in the real-time configuration of the VarEPS for weather phenomena whose initial conditions are the principal driver of its predictability. For tropical cyclone track movement, it is unclear whether this difference in initial conditions is important for real-time track bias correction since the ultimate movement of a tropical cyclone is first-order guided by the large-scale environmental steering flow (Chan and Gray 1982, Holland 1983). However this issue may be more important when utilizing the ECMWF hindcasts for real-time tropical cyclone intensity calibration.

Prior to discussing how track biases are incorporated into the real-time calibration scheme, the derivation of along-track and cross-track errors is described. Tropical cyclone track error may be decomposed into its orthogonal vector components, which are known as along-track and cross-track errors (Tsui and Miller 1988; Figure 5.1). Along-track errors are often synonymous with model timing errors in terms of track movement, whereas cross-track errors are associated with systematic biases in the directional large-scale steering flow. As a point of reference, the precedent in the tropical cyclone community is to compute along-track and cross-track errors relative to the observed track and observed heading of a tropical cyclone (Tsui and Miller 1988). However, since our proposed track correction is to use the along-track and cross-track errors derived from the model hindcasts to statistically calibrate the real-time track forecasts, the original methodology is adapted to calculate track error statistics relative to the observed track of the tropical cyclone but with respect to the forecast heading of the TC. In this respect, the derived along-track and cross-track errors may readily be used operationally, since the observed heading of a tropical cyclone represents posterior information not available at the time of forecast issuance. Note: The sign convention here is that positive along-

track error implies the model forecasts for TC track movement are *too slow* relative to observations. Positive cross-track errors indicate the forecast position is *too far west* relative to observations for a northward-moving tropical cyclone. To determine the along-track and cross-track errors as a function of forecast lead-time, the ECMWF hindcasts produced weekly for the period 2010 and 2011 during the months of June to November are utilized. As discussed in Section 2.2, the ECMWF hindcasts are a 5-member ensemble that are integrated at the same day as the current monthly forecast for each year during the previous 18-years.

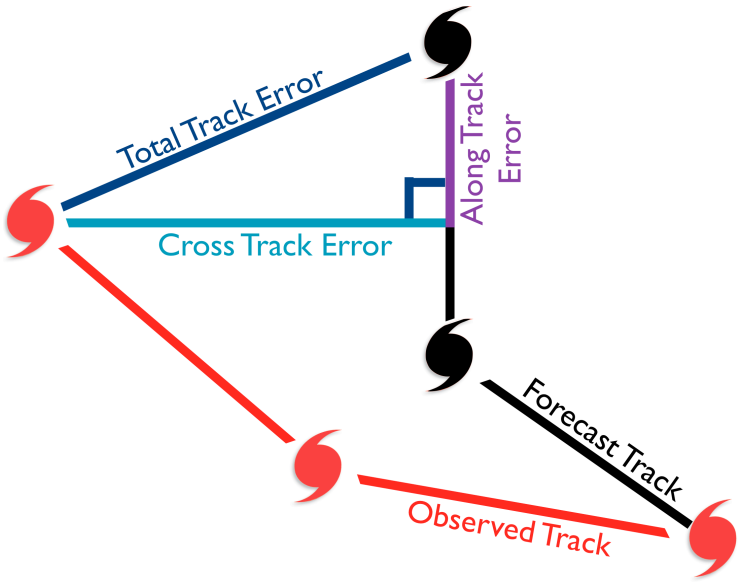

Figure 5.1: Schematic of cross-track and along-track error components for an observed tropical cyclone track (red) and forecast track (black). In this analysis, cross-track (light blue) and along-track (purple) errors are calculated with respect to the observed tropical cyclone position and the forecast track heading. In terms of the sign notation, the schematic of Figure 5.1 would result in *negative* along-track error and *negative* cross-track error.

HURDAT is used to isolate all tropical cyclones that are present in the initial conditions of the hindcast integration period 1992–2010. Next, each hindcast ensemble member is processed for tropical cyclones that occur in the initial conditions and are present within the hindcast integration period using a modified-version of the Suzuki-Parker tracking scheme (Holland et al. 2010; See Section 2.3.1 for more details). Then,

along-track, cross-track, maximum wind speed, and minimum mean sea level pressure errors are calculated as a function of forecast lead-time using the forecast locations and observed information as listed in HURDAT. To calculate the cross-track and along-track error, first a great-circle distance, $D_{i,f}$, between the initial forecast location, (x_i, y_i), and final forecast location, (x_f, y_f), is determined via Equation 5.1.

$$D_{i,f} = 2\sin^{-1}\left(\left[\sin^2\left(\frac{y_i - y_f}{2}\right) + \cos(y_i)\cos(y_f)\sin^2\left(\frac{x_i - x_f}{2}\right)\right]^{1/2}\right) \qquad 5.1$$

Next, an angular difference between the initial and final forecast latitudes, $\Delta\phi_{i,f}$, is calculated via Equation 5.2.

$$\Delta\phi_{i,f} = \log\left(\tan\left(\frac{y_f}{2} + \frac{\pi}{4}\right) \Big/ \tan\left(\frac{y_i}{2} + \frac{\pi}{4}\right)\right) \qquad 5.2$$

In Equation 5.3, the angular difference between the two forecast latitudes is used to derive the forecast course, $\delta_{i,f}$, between the initial and final forecast longitudes.

$$\delta_{i,f} = \text{atan2}\left(x_f - x_i, \Delta\phi_{i,f}\right) \qquad 5.3$$

A similar calculation using Equations 5.1–5.3 is used to derive the great-circle distance, $D_{i,o}$, angular difference, $\Delta\phi_{i,o}$, and course, $\delta_{i,o}$, between the initial forecast location and the final observed location, (x_o, y_o). These values are then used to determine the cross-track error, X, via Equation 5.4.

$$X = R\sin^{-1}\left(\sin\left(\frac{D_{i,o}}{R}\right)\sin\left(\delta_{i,o} - \delta_{i,f}\right)\right) \qquad 5.4$$

Using the geometric relationships among the cross-track error, X, the great-circle distance between the initial forecast location and final observed location, $D_{i,o}$, and the great-circle distance between the initial and final forecast locations, $D_{i,f}$, the along-track error, A, is provided in Equation 5.5

$$A = R\cos^{-1}\left(\cos\left(\frac{D_{i,o}}{R}\right)\Big/\cos\left(\frac{X}{R}\right)\right) - D_{i,f} \qquad 5.5$$

where R is the radius of earth (6371 km). After calculating the along-track, cross-track, maximum wind speed, and minimum mean sea level pressure errors for each tropical cyclone present in the initial conditions of the ECMWF hindcasts for 2010–2011, these errors are then stored as a function of forecast lead-time. It is from this historical forecast error distribution that the calibration procedure samples from to derive the corrected forecast track location.

Using a specified cross-track and along-track error sampled from the historical forecast distribution, the bias-corrected position for the forecast track of a given ensemble member is found through the following steps. First, the forecast track course, $\delta_{i,f}$, is calculated using the previous time step and position as well as the current time step and position (Equation 5.3). Next, the specified along-track error, A, along with the forecast course is added to the current latitudinal position (Equation 5.6).

$$y_{n,i} = y_f + \frac{A}{R}\cos\left(\delta_{i,f}\right) \qquad 5.6$$

The new longitudinal position from the along-track bias is found via Equation 5.7 and uses the latitudinal position found with Equation 5.6, the angular difference, $\Delta\phi_{f,ni}$ between the final forecast position and the new latitude, and the along-track error, A.

$$x_{n,i} = \mathrm{mod}\left(x_{n,f} + \frac{A}{R}\frac{\sin\left(\delta_{i,f}\right)}{\left(y_{n,i} - y_f\right)\Big/\Delta\phi_{f,n_i}} + \pi, 2\pi\right) - \pi \qquad 5.7$$

A similar procedure is used to incorporate the cross-track bias, X, onto the along-track adjusted position, ($x_{n,i}$, $y_{n,i}$), which formalizes the final latitudinal and longitudinal location, ($x_{n,f}$, $y_{n,f}$), for a particular forecast lead-time (Equations 5.8–5.9).

$$y_{n,f} = y_{n,i} + \frac{X}{R}\cos\left(\delta_{i,f} + \frac{\pi}{2}\right) \quad 5.8$$

$$x_{n,f} = \mathrm{mod}\left(x_{n,i} + \frac{X}{R}\frac{\sin\left(\delta_{i,f} + \pi/2\right)}{\left(y_{n,f} - y_{n,i}\right)/\Delta\phi_{n_i,n_f}} + \pi, 2\pi\right) - \pi \quad 5.9$$

The previous discussion highlights how a given forecast track in an ensemble member at a specific forecast lead-time is corrected after supplying an along-track and cross-track error. This approach forms the basis step for forecast track calibration that is utilized in the Monte Carlo resampling error technique. In this approach, initial position biases are removed from the initial time step and at all forecast lead-times of each ensemble forecast track. Next, the historical forecast track error distribution of cross-track and along-track errors is used. To increase the sample size of cross-track and along-track errors, especially at extended forecast lead-times, the historical forecast errors are binned into 12-hr intervals for the first 120 hr and then at 24-hr intervals at lead-times of 144 hr to 192 hr. For forecast lead-times greater than 192 hr, the along-track and cross-track errors for the 192-hr bin are used. Then, for a given ensemble forecast track, a set of 20 synthetic forecast tracks are constructed at each forecast lead-time by a random draw from the historical along-track and cross-track error distributions. The Monte Carlo resampling technique results in a set of 1,020 synthetic forecast tracks for a given tropical cyclone. Using this set of tracks for each ensemble, the median latitude and longitude is determined which in turn becomes the final calibrated ensemble forecast track position at each forecast lead-time.

The calibrated forecast tracks may then be displayed in spaghetti-track form in which all ensemble forecast tracks are overlaid. Additionally, calibrated forecast track probability information may also be derived, whereby forecast track probabilities are calculated on a regular latitude-longitude grid with horizontal resolution of 0.25^{o} and use a one-degree distance threshold to determine track membership. The final track

probabilities are then normalized with respect to the total number of synthetic forecast tracks.

5.2.2. Tropical Cyclone Intensity Calibration

The tropical cyclone intensity forecasts are calibrated using a quantile-to-quantile (q-to-q) mapping technique (e.g. Calheiros and Zawadzki 1987, Webster et al. 2006, Hopson and Webster 2010) to ensure that the forecast model's climatological frequency of some variable at each quantile agrees with its observed climatological frequency. The q-to-q technique preserves the model's spatial and temporal covariances but removes the forecast bias at each quantile nonlinearly. Here, the q-to-q approach is adapted to the tropical cyclone intensity problem by mapping the ECMWF hindcast intensities for the period 1992–2010 to the observed maximum wind speeds of tropical cyclones listed in HURDAT for the same time period.

Although the q-to-q technique removes forecast intensity bias at each quantile, this adjustment does not necessarily increase forecast skill. In other words, if critical dynamical aspects of the tropical cyclone intensification process are missing from the VarEPS (e.g., rapid intensification), then other approaches become necessary to correct these deficiencies. However, the working hypothesis is that the q-to-q calibration approach will improve upon the raw VarEPS intensity forecasts due to the nonlinear intensity biases that have been routinely observed by the author in the North Atlantic basin and elsewhere during the period 2007–2011.

A review of the available literature provided no previous instances where the q-to-q technique has been used to calibrate model-derived intensity forecasts. Instead, the primary research thrust in improving forecasts of tropical cyclone intensity has been to run very high-resolution nested simulations of tropical cyclones, enhancing data assimilation techniques and/or increasing in situ/remote observations of the inner-core of a tropical cyclone (Rappaport et al. 2012). Ironically, in spite of over a decade of

research devoted to improving tropical cyclone intensity forecasts generated from dynamical models, to date the most skillful operational intensity forecast in the North Atlantic is a statistical approach using a rather simple logistic growth equation model (DeMaria et al. 2010, Cangialosi and Franklin 2012). It is this result that has led the author to conclude that the most likely path forward to improving intensity forecasts from dynamical models in the short-term is through statistical post-processing.

5.2.3. Developing the High-Predictable Forecast Cluster

The primary goal of the statistical-dynamical forecast approach is targeted at ensemble forecast calibration primarily to remove model bias and to increase ensemble dispersion. A key element in the improvement of tropical cyclone forecasts has been the development of an ensemble clustering approach that increases the tropical cyclone track prediction skill in situations with large forecast uncertainty. Whereas other groups such as ECMWF utilize clustering techniques to categorize the principal large-scale flow-regimes present within the VarEPS (Ferranti and Corti 2011), here clustering is used to isolate the ensembles within the VarEPS that are most likely to lie closest to reality throughout the forecast integration. In this approach, the top five ensemble members of the VarEPS are identified whose average correlation coefficient with the ECMWF high-resolution deterministic forecast track through a lead-time of 72 hr is largest. During pre-operational evaluation of the forecast clustering technique, a variety of forecast windows for clustering were considered including: 24-hr, 48-hr, 72-hr, and 96-hr. Ultimately, the 72-hr window produced the high-predictable cluster with the overall lowest average track error. Another approach using the smallest great-circle distance separation between the deterministic and VarEPS ensembles was also considered. However, the covariance-based technique produced the most robust cluster with the smallest total track error. The rationale behind this approach is that on average the ECMWF deterministic provides a superior track and intensity forecast to the mean VarEPS through this lead-time.

However the ensemble mean or maximum likelihood of the ensembles is expected to provide a superior forecast at extended lead-times. In the case of maximizing tropical cyclone track predictability in cases of large ensemble spread, the hypothesis is that if a subset of ensembles can be identified that provide a more accurate track forecast in the short-term than the full VarEPS, then this cluster is more likely to provide a superior track forecast with a smaller cone of uncertainty than the full VarEPS at extended forecast lead-times as well. This working hypothesis is evaluated in Section 5.3.3 for the years 2009–2011.

5.2.4. Intraseasonal Tropical Cyclone Forecasts

In Chapter 4, the ECMFS is shown to produce skillful tropical cyclone forecasts on intraseasonal time scales during the evaluation period of 2008–2009 for portions of the southern Gulf of Mexico, Caribbean Sea and the MDR. From these findings, the design of the operational forecast system for intraseasonal time scales is now discussed which utilizes the twice-weekly ECMFS, weekly hindcast product, the historical climatology of tropical cyclone tracks in the North Atlantic, and the phase and amplitude of the MJO.

In designing the operational forecast system for the intraseasonal period, there are several post-processing steps necessary to correct for systematic model biases in order to generate a more skillful forecast. Figure 5.2 is a flowchart for the process that has been designed thus far and is operational twice-weekly on Mondays and Thursdays. The first step is to apply the tropical cyclone tracking scheme on the ECMFS and equivalent hindcast products from the latest dissemination (see Section 2.3.2 for more information on the tracking scheme). In addition, tropical cyclone tracks are determined for the 5-member hindcast that is integrated from the same day and month of the year for each of the last 18 years. Next, the observed location of all tropical cyclones during the hindcast integration period is extracted from HURDAT. This set of observed tracks forms the

basis dataset for developing the observed tropical cyclone track climatology for the latest intraseasonal forecast. Afterwards, the hindcast tracks that originate in the MDR of the tropical Atlantic are compared to the observed TC tracks that also originate in this region. During the peak hurricane season, the genesis location of TCs in this region from the hindcasts is systematically shifted northward and eastward relative to the observed climatology. This bias reflects northward model drift in the preferred track of AEWs in association with the mid-tropospheric extension of the African easterly jet into the tropical Atlantic. The northward bias is thought to be due to a stronger than normal cross-equatorial temperature gradient between the Gulf of Guinea and Saharan heat low driven by a cold bias in sea surface temperatures in the Gulf of Guinea (Frederick Vitart, personal communication). To render the ECMFS tracks more useful, a mean genesis latitude and longitude is determined between TCs in the hindcast period and the observed climatology. Then, this correction factor is applied to all TCs during the hindcast integration period along with TCs from the real-time ECMFS that originate in the same region. The real-time TC intensity forecasts are adjusted using a quantile-to-quantile intensity correction derived from the 15-day TC forecasts (see Section 5.2.2).

Next, the observed and bias-corrected hindcast TC tracks are translated onto a regular 2.5^{o} x 2.5^{o} latitude-longitude grid for the tropical Atlantic. As discussed in Chapter 4, a coarse resolution at these extended time scales is utilized since the primary goal is to characterize regional changes in the forecast track density distribution. The observed climatology for a location is defined as the number of unique tropical cyclones that are within 2.5^{o} of each grid point, normalized by the total number of years of the climatology. Since the observed climatology is compared to the model-derived climatology, the normalization factor is 18 years. The model-derived climatology is calculated in an equivalent way: for each location the number of unique ensemble TC tracks that are within 2.5^{o} is determined, and this total is normalized by 90 (18 total years with 5 ensembles per year). To derive the real-time track density distribution, a similar

procedure is utilized by counting the number of unique ensemble TC tracks within 2.5° but normalization is with respect to the total ensemble size (51). Next, local differences at each grid point between the observed and model climate become bias-adjustment factors that are then added or removed from the real-time track density distribution. In addition, track density anomalies are also calculated except that they are with respect to the model hindcast climatology.

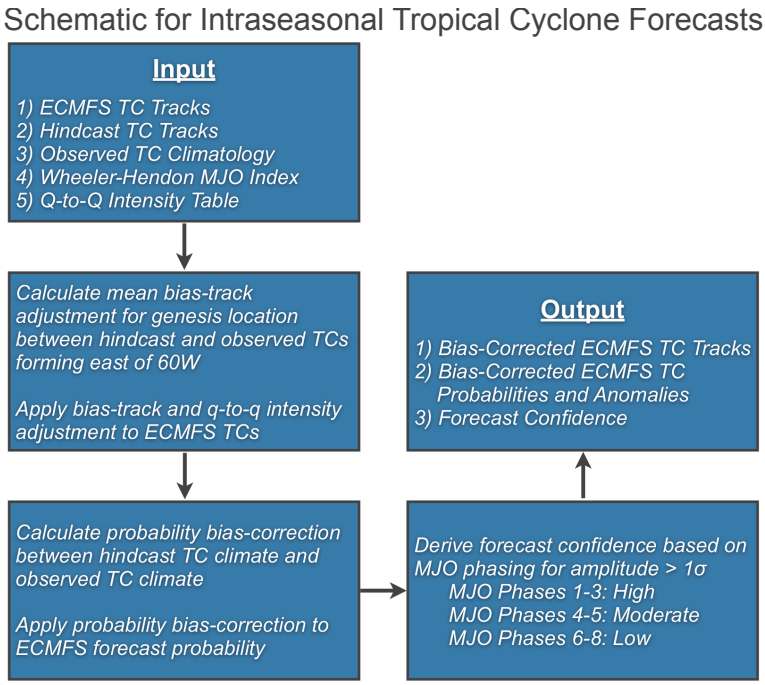

Figure 5.2: Background methodology on how the intraseasonal tropical cyclone forecasts for the North Atlantic are constructed from the ECMWF Monthly Forecast System.

5.3. Results

To evaluate the forecast calibration procedures that have been incorporated into the operational forecast system for North Atlantic TCs, first the distribution of track and intensity errors of the hindcasts for 2010–2011 are reviewed in Section 5.3.1, as these errors are utilized in the real-time track and intensity calibration scheme. Then, in Section 5.3.2, the uncorrected and calibrated ECMWF deterministic and VarEPS track forecasts for 2010–2011 are compared to the NHC performance for the same time period.

These findings are followed by an evaluation of the forecast track performance of the high-predictable cluster relative to the calibrated deterministic and mean VarEPS track forecasts in Section 5.3.3. Then in Section 5.3.4 the calibrated intensity forecasts are verified using Saffir-Simpson intensity categories and the ranked probability score. Finally, the uncorrected and calibrated intraseasonal forecasts for the period 2008–2011 are evaluated in Section 5.3.5.

5.3.1. ECMWF Hindcast Track and Intensity Errors

Figure 5.3a (Figure 5.3b) shows the along-track (cross-track) errors as a function of forecast lead-time from the ECMWF hindcasts that were initialized during the period May to November of 2010–2011, after correcting for initial position errors of the initialized tropical cyclone. Both orthogonal error components show Gaussian distributions with small but nonzero biases during the first 96 hr that systematically increase toward larger errors as forecast lead-time increases. In addition, the ECMWF hindcasts show significantly larger cross-track errors relative to along-track errors at large lead-times. For instance, by 120 hr (192 hr), the median along-track error is 21 n mi (48 n mi) whereas the median cross-track error is 61 n mi (121 n mi). The positive along-track and cross-track biases at extended forecast lead-times indicate that hindcast TCs tend to move too slow and are located too far to the west relative to a northward-moving TC.

Figure 5.4a shows the maximum wind error from the ECMWF hindcasts as a function of forecast lead-time after removing initial intensity errors. No significant systematic biases with forecast lead-time are found in the ECMWF hindcasts after adjusting for initial intensity errors, as the wind errors are normally distributed about zero. However, when the relationship in maximum wind speed between observations and the ECMWF hindcasts is viewed using a q-to-q technique, systematic nonlinear biases are seen that also vary with horizontal resolution (Figure 5.4b). First, for the tropical

depression and tropical storm stage, the high-resolution ECMWF hindcasts (i.e. 0.25^{o} for forecast days ≤ 10) exhibit a small negative intensity bias in which the ensembles are climatologically too weak relative to observations. Then, in the intensity range of a Category one hurricane, the hindcasts show no bias relative to observations. Once maximum winds exceed Category one stage, significant positive biases emerge. This result for the high-flow regime indicates that climatologically, the ECMWF hindcasts produce too frequently intense cyclonic vortices relative to their observed frequency in the hindcast integration period. For the ECMWF hindcasts at the coarser 0.5^{o} resolution (i.e. forecast days > 10), the q-to-q map is transformed such that the biases exhibited during the high-resolution integration are amplified. For instance, in the tropical storm through Category 2 hurricane range, the ECMWF hindcasts are significantly weaker relative to observations with the maximum magnitude of the negative wind bias of about 15 kts (7.7 m s^{-1}).

Assuming these track and intensity errors are representative of the biases found in the real-time configuration of the VarEPS, they provide an opportunity to statistically post-process the VarEPS to ensure the corresponding track and intensity forecasts are well-calibrated relative to observations. However, one caveat is that unlike tropical cyclone track forecasts, which are primarily guided by the large-scale steering flow, the tropical cyclone's forecast intensity is strongly dependent on the initial representation of the cyclonic vortex and its 3D structure in the model analysis fields. Since the initial conditions of the ECMWF hindcasts are derived from the ERA-Interim reanalysis, which like ERA-40 has a coarse representation of intense cyclonic vortices that are not 'bogused' but are assimilated based on observations (e.g. Sriver and Huber 2006), the assumption is that the hindcast intensity error distributions even after accounting for initial intensity biases may not be representative of real-time intensity errors in the VarEPS. This expectation is to due to the varying methods in how initial conditions are generated between the frozen configuration of the ECMWF atmospheric model that is

used in the ERA-Interim reanalysis and the more recent version of the atmospheric model that features an ensemble of data assimilation technique to supplement the weak-constraint 4D-Var approach. In addition, the initial conditions of the real-time VarEPS are cast on a high-resolution grid at a nominal resolution of 30 km, whereas the initial conditions in the ECMWF hindcasts use interpolated fields of the ERA-Interim reanalysis which are stored at a nominal resolution of 80 km.

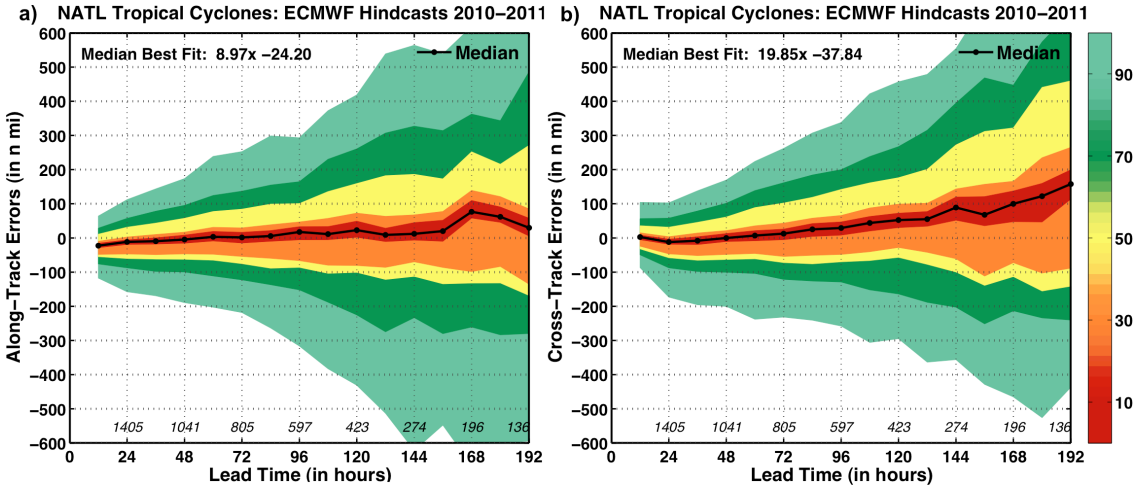

Figure 5.3: a) Along-track and b) cross-track errors expressed as interpercentiles for all tropical cyclones present in the initial conditions of the ECMWF hindcasts initialized during May to November for 2010 and 2011.

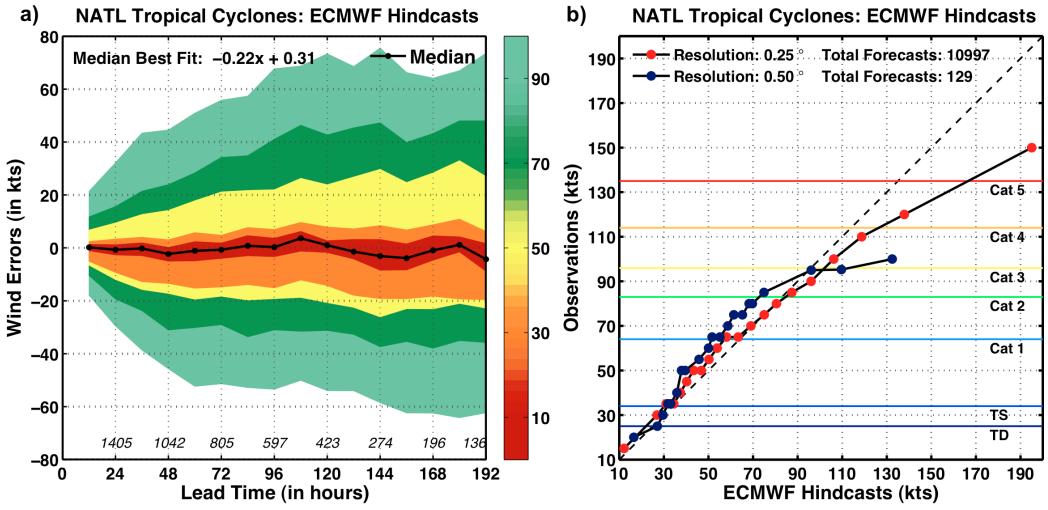

Figure 5.4: a) Similar to Figure 5.3a, except for wind errors. b) Quantile-to-quantile scatter plot of maximum intensity forecasts derived from ECMWF hindcasts and observed maximum intensity from HURDAT for the period 1992–2010.

5.3.2. VarEPS Track Forecast Performance during 2010–2011

To demonstrate the impact that forecast track calibration has on the deterministic and VarEPS forecasts, Table 5.1 lists the homogenous total track error (in n mi) for forecasts from 2011 as a function of lead-time for the uncorrected and calibrated ECMWF deterministic and mean VarEPS. For comparison purposes, homogeneous track errors for the NHC are also included. Overall, the Monte Carlo resampling error technique in conjunction with the initial position bias correction leads to the most significant improvement at short lead-times, with the relative improvement decreasing as forecast lead-time increases. During the first 48 hr, the track errors with the calibrated deterministic and mean VarEPS are reduced by 39% and 33%, respectively. For lead-times of 48–96 hr, the average improvement in the track forecasts for the calibrated deterministic and mean VarEPS is 22% and 14%, respectively. However, at lead-times of 120 hr and beyond, the forecast track calibration scheme does not improve upon the uncorrected deterministic, although there is a small nominal improvement in the mean VarEPS. When compared to the official forecasts released by the NHC, the uncorrected ECMWF deterministic has lower track errors at lead-times of 96 hr and beyond, whereas the calibrated deterministic has lower track errors at lead-times of 72 hr and beyond. Relative to the uncorrected and calibrated mean VarEPS, the NHC track errors are lower at all forecast lead-times.

Similar to Table 5.1, homogeneous track error statistics are presented in Table 5.2 for the 2010 hurricane season. For the first 48 hr, the average relative improvement in track forecasts between the uncorrected and calibrated ECMWF deterministic and mean VarEPS is 33% and 23%, respectively. For lead-times of 48–96 hr, the relative improvement decreases to 12% and 3%, and at extended lead-times of 120 hr and beyond, the forecast track correction scheme does not yield significant improvement for either the deterministic or the mean VarEPS. In fact, the calibrated deterministic and mean VarEPS actually show a slight degradation in track forecast performance for

extended lead-times. The lack of track error improvement beyond 120 hr for both 2010 and 2011 seasons is attributed to systematic differences in the VarEPS's distribution of along-track and cross-track biases at these lead-times relative to historical track biases from the ECMWF hindcasts. When these forecasts are compared to the NHC from 2010, overall the NHC's track forecasts are superior relative to the uncorrected and calibrated deterministic and mean VarEPS. For lead-times of 48 hr to 96 hr, the calibrated deterministic is competitive with the NHC's track performance but by 120 hr, the NHC has significantly lower track errors. In fact, the 187 n mi error at 120 hr is well-below the NHC's typical performance at this lead-time (213 n mi over the last five years; Cangialosi and Franklin 2011).

Table 5.1: Mean homogeneous total track error (in n mi) as a function of forecast lead-time for the NHC, ECMWF Deterministic, calibrated ECMWF Deterministic, mean VarEPS, and calibrated mean VarEPS for 2011. The values listed in brackets denote the 50% interval about the median error (in n mi). The values in bold are lower than the respective uncorrected deterministic or mean VarEPS.

	NHC	Deterministic	Calibrated Deterministic	Mean VarEPS	Calibrated Mean VarEPS
12 hrs	28 [13–37]	92 [28–127]	**52 [28–71]**	95 [34–130]	**54 [31–70]**
24 hrs	44 [24–60]	104 [43–141]	**61 [27–78]**	108 [49–145]	**70 [37–91]**
36 hrs	59 [30–78]	115 [55–150]	**76 [37–98]**	125 [61–160]	**88 [49–114]**
48 hrs	73 [36–110]	120 [57–171]	**81 [39–105]**	136 [74–177]	**106 [57–140]**
72 hrs	113 [53–165]	133 [67–172]	**103 [55–137]**	171 [87–222]	**155 [83–197]**
96 hrs	168 [93–230]	165 [83–208]	**149 [101–179]**	227 [109–279]	**216 [114–268]**
120 hrs	244 [137–366]	213 [108–310]	225 [147–258]	306 [143–381]	**303 [158–389]**
144 hrs	N/A	307 [157–383]	338 [197–448]	389 [186–494]	390 [204–519]
168 hrs	N/A	417 [214–536]	423 [229–535]	476 [240–635]	**463 [253–615]**

Table 5.2: Similar to Table 5.1, except for 2010.

	NHC	Deterministic	Calibrated Deterministic	Mean VarEPS	Calibrated Mean VarEPS
12 hrs	33 [17–45]	96 [33–143]	**57 [30–67]**	101 [40–144]	**65 [34–81]**
24 hrs	52 [29–71]	100 [40–137]	**66 [34–88]**	114 [54–154]	**84 [44–107]**
36 hrs	71 [37–94]	115 [54–150]	**81 [37–108]**	132 [70–172]	**108 [58–139]**
48 hrs	89 [51–81]	118 [60–159]	**89 [45–122]**	146 [81–190]	**129 [71–171]**
72 hrs	131 [81–159]	147 [69–199]	**134 [58–202]**	189 [100–247]	**183 [98–244]**
96 hrs	168 [94–199]	178 [96–249]	**174 [77–243]**	225 [117–293]	236 [128–304]
120 hrs	187 [97–221]	220 [115–294]	228 [115–330]	269 [138–269]	287 [153–375]
144 hrs	N/A	259 [129–332]	269 [154–367]	321 [159–419]	347 [191–442]
168 hrs	N/A	306[184–394]	352 [223–472]	391 [192–504]	454 [245–581]

Previously, the change in track error is considered using the uncorrected and calibrated ECMWF deterministic and mean VarEPS forecasts. A probabilistic approach to track verification is now utilized to consider how the conditional bias from the track forecasts changes as a function of forecast probability level and lead-time. The reliability diagrams in Figure 5.5 show the relationship between forecast probability level and observation frequency at forecast days ≤ 2, 2–4, 4–7, 7–10, and 10–15. For the purpose of this comparison, the variability in these statistical benchmarks is also considered relative to increasing interpercentile ranges at intervals of 10%, 25%, 50%, 75%, and 95% with the Monte Carlo resampling technique. These intervals are selected in order to determine the optimum window for probabilistic forecast calibration. Note: The 95% resampling interval is used operationally in the tropical cyclone track calibration scheme. The salient features from Figure 5.5 include the following. Regardless of the resampling interval specified, the calibrated forecast track probabilities show similar reliability to the uncorrected VarEPS, and any differences in the conditional bias are the result of deviations in the relative frequency of the number of forecasts for a particular probability

level (Figure 5.5 insets). Since the calibrated VarEPS forecasts do not lie on the one-to-one diagonal, this suggests that additional post-processing techniques could be utilized to increase their reliability. The maximum forecast probability level (observation frequency) decreases as forecast lead-time increases. Specifically, these values are: 100% (95%) for forecast days ≤ 2, 50–60% (35–50%) for forecast days 2–4, 40% (25%) for forecast days 4–7, and 20% (10%) beyond 7 days. This information is important for users who have varying degrees of risk tolerance. For instance, end-users who require a high degree of confidence before using a probabilistic track forecast will only be able to utilize the calibrated VarEPS track probabilities through a lead-time of about four days, whereas those users who can tolerate larger uncertainty should find utility in the probabilistic forecasts through seven days.

In comparison to the uncorrected VarEPS forecasts, Figure 5.5 shows no substantial difference in reliability with the calibrated VarEPS probability forecasts. This result suggests the distinguishing factor in whether the track probability forecasts are improved by track calibration is if they provide greater statistical resolution. Table 5.3 summarizes the change in resolution and forecast skill of the uncorrected and calibrated VarEPS track probability forecasts using the ROC score. ROC scores are above 0.5 at all forecast lead-times indicating that the uncorrected and calibrated VarEPS track forecasts provide useful forecast information through two weeks in advance. At all forecast lead-times, ROC scores increase as the interpercentile range in the Monte Carlo resampling technique increases. Relative to the ROC scores for the uncorrected VarEPS track forecasts, the calibrated forecasts have larger ROC scores, so long as the resampling interval is 25% or larger. Finally, the calibrated track approach using a 95% resampling interval, which is used in the operational forecast scheme, increases the ROC score on average by 8% compared to the uncorrected VarEPS, with the most significant improvement found during forecast days 4–7 and 7–10 where the ROC scores are increased by 10% and 17%, respectively.

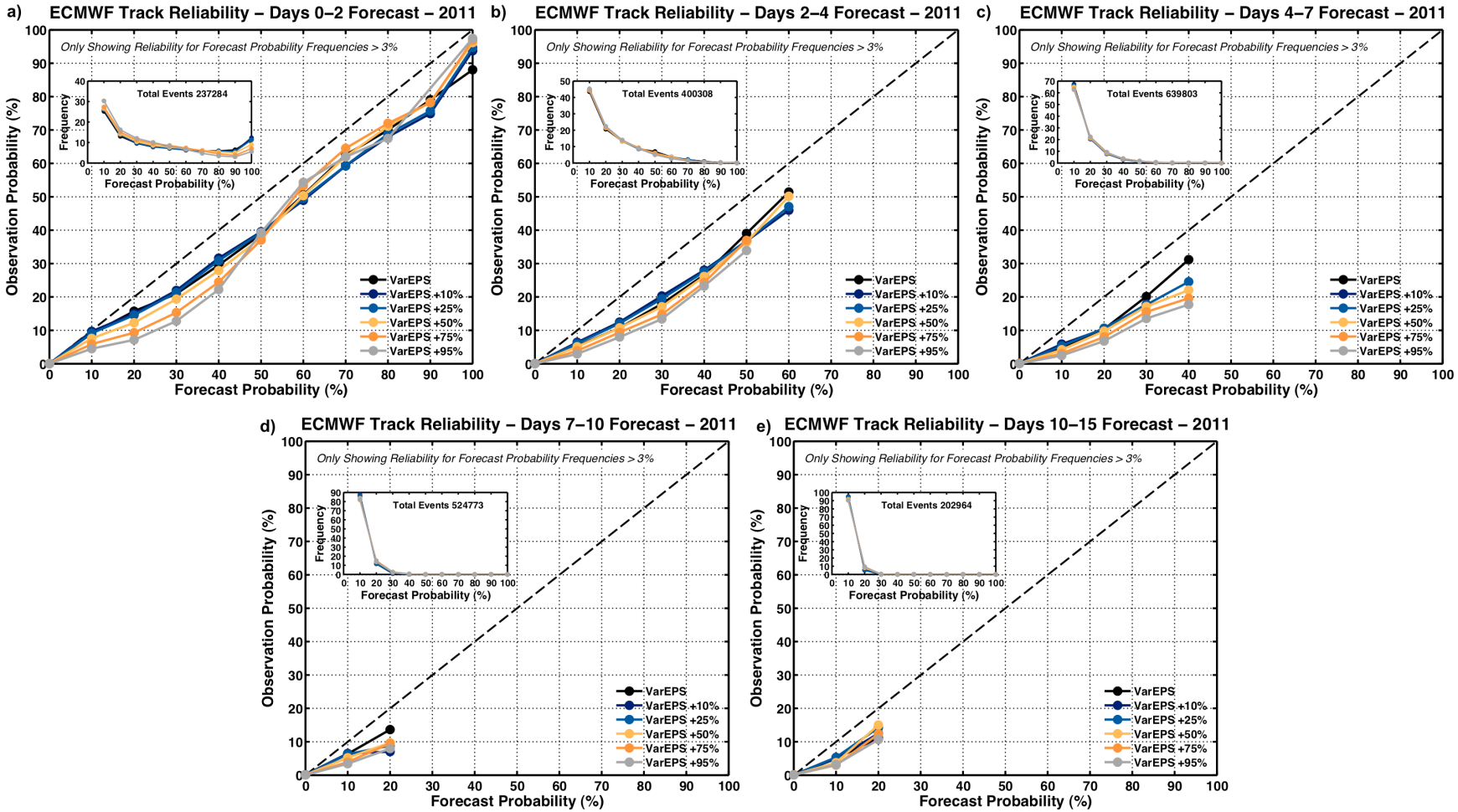

Figure 5.5: Reliability diagrams of the ECMWF VarEPS for forecast days a) ≤ 2, b) 2–4, c) 4–7, d) 7–10, and e) 10–15.

Table 5.3: Relative operating characteristic (ROC) score of the ECMWF VarEPS forecasts for 2011. Columns with +XX% denote ROC scores for increasing resampling interpercentile ranges about the median along-track and cross-track error distributions derived from the ECMWF hindcasts.

ROC Score 2011	ECMWF VarEPS	ECMWF VarEPS +10%	ECMWF VarEPS +25%	ECMWF VarEPS +50%	ECMWF VarEPS +75%	ECMWF VarEPS +95%
Days ≤ 2	0.95	0.95	0.95	0.96	0.96	0.96
Days 2–4	0.91	0.91	0.91	0.93	0.94	0.94
Days 4–7	0.79	0.77	0.79	0.83	0.86	0.87
Days 7–10	0.63	0.59	0.62	0.69	0.73	0.73
Days 10–15	0.56	0.57	0.58	0.60	0.61	0.60

5.3.3. Track Forecasts with the High-Predictable Ensemble Cluster

Although the primary goal in designing the operational forecast system is to improve deterministic and probabilistic track forecasts through calibration, one of the reoccurring challenges in tropical cyclone track forecasting and with weather forecasts in general is how to handle a forecast that has large uncertainty. Many operational agencies and forecast centers utilize the mean of the ensemble distribution as an indication of the most likely outcome in these situations with limited predictability. Often, however, the ensemble distribution for forecasts with large uncertainty will not be normally distributed and instead will showcase a variety of modal solutions.

One approach that is utilized here samples from the full distribution of VarEPS ensembles to select those members that are more likely to align closer to observations. In this case, a high-predictable clustering technique has been developed to take advantage of the superior performance of the ECMWF high-resolution deterministic forecast during the short-term along with the full set of VarEPS ensembles. Using the ideas from Palmer and Anderson (1994) for subseasonal and seasonal forecasts, the VarEPS ensembles that

deviate significantly from reality during the first few days are less likely to represent the observed outcome at later time steps. The analogy to the tropical cyclone clustering routine is that on average the ECMWF deterministic provides a superior track forecast than the mean VarEPS during the first several days of the model integration (i.e., Section 5.3.2). Therefore, if the deterministic forecast is treated as 'reality' for the first few days of the model integration, then a subset of VarEPS ensembles that align most closely to the deterministic solution may be selected. The working hypothesis is that the resulting cluster is more likely to align closer to observations than either the deterministic or mean VarEPS solutions especially with forecasts that feature large uncertainty. This idea is evaluated systematically using all deterministic and VarEPS post-genesis track forecasts issued during the period 2009–2011. The high-predictable ensemble cluster is created by selecting the top five VarEPS ensembles whose average latitudinal and longitudinal correlation coefficient with the ECMWF deterministic track forecast is largest during the first 72 hr.

Figure 5.6a shows the total track error as a function of forecast lead-time for the high-predictable cluster, calibrated deterministic, and calibrated mean VarEPS during the period 2009–2011. Compared to the ensemble mean, the high-predictable cluster on average results in track errors that are lower by 12% during the first six days. However, even at lead-times from three to six days, the calibrated ECMWF deterministic on average has significantly lower track errors than the ensemble mean (by 25%) or the high-predictable cluster (by 14%). Although these results are valid for all forecasts, a different picture emerges if the forecasts are stratified by ensemble spread, defined here as the standard deviation of the VarEPS. In this case, ensemble spread is used as an a priori indicator of forecast error or uncertainty. When the ensemble spread of the VarEPS is larger (smaller) than normal, the mean VarEPS and ECMWF deterministic yields larger (smaller) forecast track error relative to normal. When the high-predictable cluster is compared to the ECMWF deterministic using forecasts when the ensemble

spread is larger than average, Figure 5.6a shows that beyond a lead-time of 108 hr, the high-predictable cluster provides a track forecast with the smallest overall error and is on average 10% smaller. In addition, although the deterministic forecast provides a superior forecast when compared to the ensemble mean even when stratified by large ensemble spread, the relative difference is much smaller (by 10%) than if all forecasts are considered.

The high-predictable cluster also leads to a narrowing of the ensemble cone-of-uncertainty as seen using the standard deviation of the ensemble track error (Figure 5.6b). For instance, using the ensemble standard deviation to define the ensemble spread, the average spread in the VarEPS at 72 hr is 83 n mi, whereas the high-predictable cluster is only 60 n mi. By 120 hr, the average ensemble spread of the VarEPS grows to 150 n mi, while the high-predictable cluster is within 100 n mi. On average, the spread in the ensemble cluster is 27% smaller than the full VarEPS.

Although the high-predictable cluster results in substantially smaller spread than the full VarEPS, the spread–skill relationship is not as robust. Using the correlation coefficient between the ensemble standard deviation and mean VarEPS track error to define the spread–skill relationship, the average correlation coefficient for the first 144 hr is 0.65 for the VarEPS, while the high-predictable cluster is only 0.31. In other words, while the high-predictable cluster yields track errors that are substantially lower than the mean VarEPS, the reduction in track error is not large enough to counteract the significant narrowing in ensemble spread. Thus, a more appropriate way to view the track forecasts from the high-predictable cluster is that it offers a complementary solution to the mean VarEPS and deterministic track forecasts and will on average provide a better track forecast than the other two solutions at extended lead-times when the VarEPS spread is larger than normal. However, the high-predictable cluster should not be treated as a refined cone-of-uncertainty due to its weak spread–skill relationship.

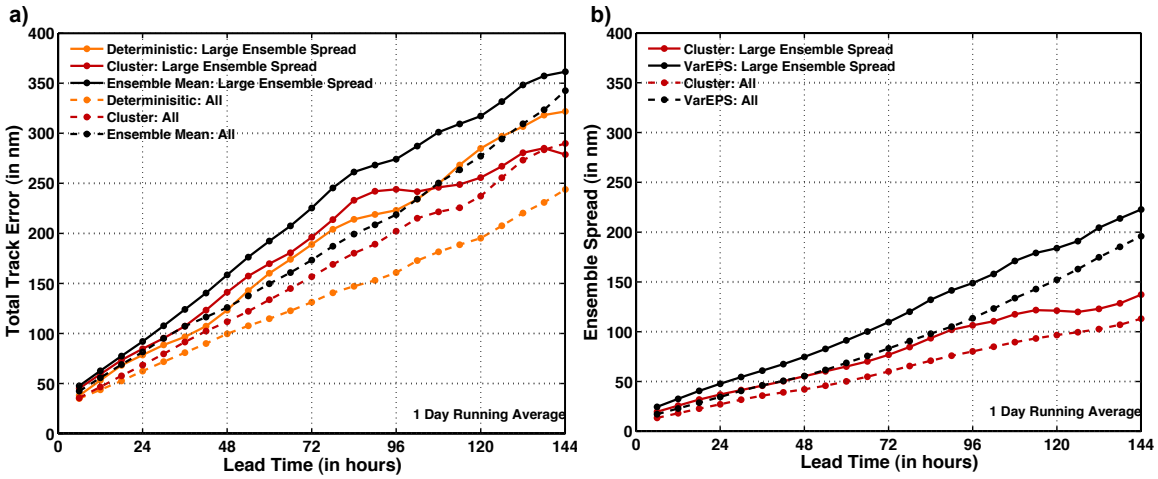

Figure 5.6: a) Time series of homogeneous total track error (in n mi) for the calibrated ECMWF deterministic (dashed orange), calibrated mean VarEPS (dashed black), and median of the high-predictable cluster (dashed red) for the period 2009–2011. b) Similar to a) except for average ensemble spread by forecast lead-time. The solid lines are for forecasts when the ensemble spread during the first 84 hr is greater than the average standard deviation of all forecasts. All time series have been filtered once using a 5-point running mean.

5.3.4. VarEPS Intensity Forecast Performance during 2010–2011

Recognizing the difficulty in forecasting tropical cyclone intensity using the maximum sustained 1-minute wind rounded to the nearest 5 kt, another approach that is often adopted is to forecast Saffir-Simpson intensity categories. When Herbert Saffir and Robert Simpson developed the original Saffir-Simpson scale, the intention was to produce a rather simple intensity index that conveyed the wind damage potential for a landfalling hurricane (Simpson 2012). Recent studies have indicated that the damage potential from a landfalling hurricane is actually more dependent on other variables than simply maximum wind speed such as tropical cyclone size, translation speed, and integrated kinetic energy (Powell and Reinhold 2007, Holland and Done 2011). However, given the saturation of the Saffir-Simpson scale in the tropical cyclone community, energy trading industry, and the general public, here the tropical cyclone intensity forecasts are evaluated by stratifying the forecasts and observations into four

categories which include: tropical depression (max winds < 35 kts), tropical storm (35 kts ≤ max winds < 64 kts), hurricane (64 kts ≤ max winds), and major hurricane (95 kts < max winds) for the years 2010 and 2011. To verify the probabilistic intensity forecasts with observations using this category system, the ranked probability score is used and is shown in Figure 5.7a. Similar to the BS, the RPS has a range of zero to one with lower scores indicative of better-forecast performance than larger scores. The best RPSs are obtained if the forecast probabilities by category are sharp and if observations fall within the category with the highest forecast probability. Note: For large sample sizes, a RPS of 0.25 indicates the skill of random chance alone for a four-category forecast system. During the first 96 hr, the RPSs increase reflecting a steady degradation in the forecast utility of the uncorrected VarEPS to capture correctly maximum TC intensity. Beyond four days and through a lead-time of eight days, the RPSs remain fairly constant around 0.25, which reflects a saturation point in terms of how large the intensity errors may grow and indicates that the uncorrected VarEPS intensity forecasts are not skillful relative to random chance alone beyond 96 hr. In addition, based on the interpercentile plumes of the forecast intensity distribution, ensemble intensity dispersion does not increase beyond the first 36 hr.

A similar evaluation of the VarEPS intensity forecasts is conducted except for the intensity forecasts that have been bias-corrected in terms of removing initial intensity errors and applying the nonparametric q-to-q calibration technique developed from the ECMWF hindcasts. Figure 5.7b shows the relative improvement in the RPS using the statistical post-processed intensity forecasts relative to the uncorrected VarEPS in Figure 5.7a. The greatest improvement is found during the first 72 hr in which the statistically-rendered forecasts have an average median improvement around 40%. However, beyond 96 hr and through 192 hr, the median improvement is about 15–20% depending on the lead-time considered. More importantly, at these forecast lead-times the RPS for the

calibrated VarEPS intensity forecasts reach a saturation limit of 0.2 indicating these forecasts have skill beyond random chance alone through a lead-time of 192 hr.

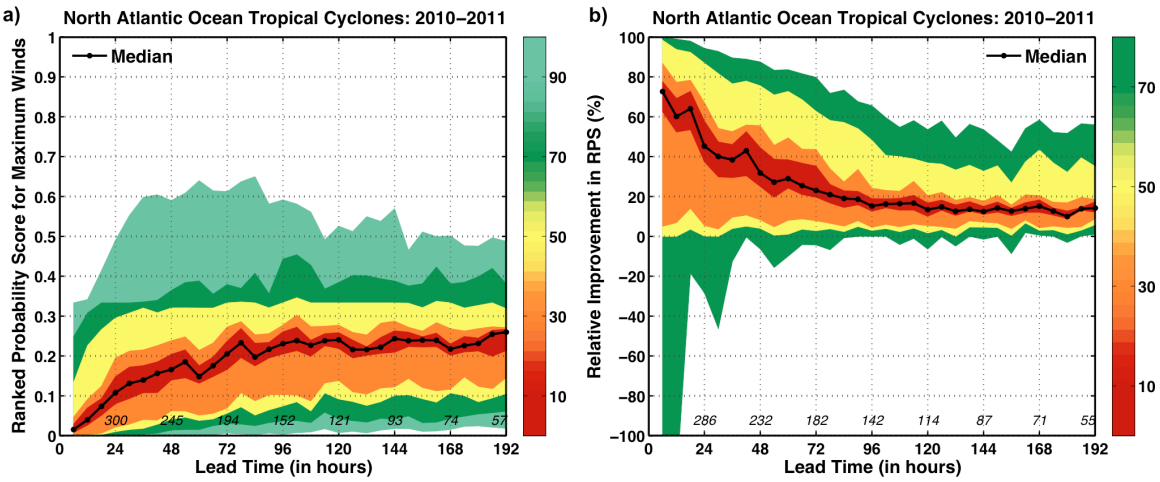

Figure 5.7: a) RPSs of the uncorrected ECMWF VarEPS tropical cyclone intensity forecasts for 2010 and 2011 using Saffir-Simpson categories of tropical depression, tropical storm, hurricane, and major hurricane. b) Relative improvement (in %) in RPS of the ECMWF VarEPS after incorporating initial intensity adjustment and quantile-to-quantile intensity calibration.

5.3.5. Calibrated Intraseasonal Tropical Cyclone Forecasts

Here, the calibrated intraseasonal forecasts for the period 2008–2011 is considered, first with an example of the impact that statistical-post processing of the ECMFS has on the intraseasonal forecasts. Figure 5.8a (Figure 5.8b) shows the uncorrected (calibrated) TC track density probability and anomaly computed with respect to the observed (model) climatology for the forecast initialized on 5 August 2010 for the period 16 August – 6 September 2010. The observed track of tropical cyclones during the forecast period is also included and shows that this period was unusually active especially in terms of the total number of tropical cyclones that moved through the Atlantic basin and with a preferred track for TCs to move north of the Greater Antilles and recurve east of the U.S. Atlantic coast. Compared to the calibrated forecast track probability in Figure 5.8b, the forecast probabilities in Figure 5.8a are underestimated in some cases by a factor of two, indicating that the ECMWF hindcasts significantly

underpredict the observed climatological distribution of tropical cyclones for the 18-year period. In addition, the underestimation of track density is spatially inhomogeneous, with the Caribbean and Gulf of Mexico showing a larger negative track density bias than the subtropical Atlantic. This result indicates the ECMWF hindcasts tend to under forecast the genesis and track of tropical cyclones farther west through the Atlantic basin. When the calibrated ECMFS track density is considered relative to the model's climatological distribution (Figure 5.8b), the regions with substantially larger track density anomalies are collocated with the observed distribution of TC tracks for this time period.

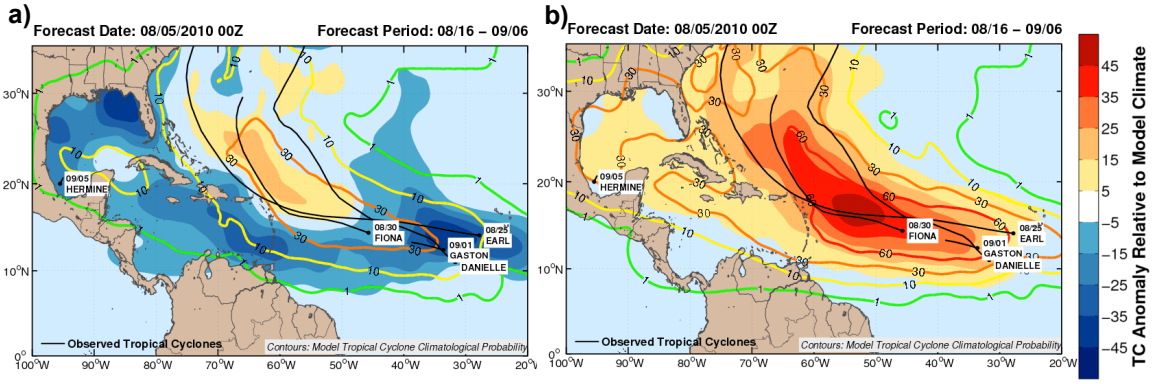

Figure 5.8: a) Example of ECMFS TC track density anomaly (shaded) relative to the observed climatology from 1992 to 2009 with contours denoting forecast probabilities for the forecast initialized on 5 August 2010. Forecast probabilities are contoured at intervals of 1%, 10%, 30%, and 60%. b) Similar to a), except track density anomalies are relative to model hindcast climatology from 1982–2009 and contours are bias-corrected forecast probabilities.

In Chapter 4 a TC predictability study was presented in which the initial phase and amplitude of the MJO is shown to modulate both the observed and uncorrected ECMFS distribution of tropical cyclone activity for the years 2008 and 2009. However, the limited sample size of unique MJO events during these two years raised caution with some of the results and the proposed ECMFS–MJO sensitivity. This relationship is now reconsidered using calibrated versions of the ECMFS from 2008 to 2011 for forecast days 11–32. Figure 5.9a shows the forecast track density anomaly of the calibrated ECMFS relative to the ECMWF model climate for MJO Phases 1-3 along with the

calibrated forecast probabilities. The calibrated forecast track density anomalies are 10–15% above normal for the MDR stretching northward of the Greater Antilles and east of the United States which agrees well with the distribution of observed TC tracks for the same time period. In addition, maximum forecast probabilities in the MDR reach a peak value in the moderate range of 30–60%. Relative to the other phase groups of the MJO (Figure 5.9b – MJO Phases 4-5; Figure 5.9c – MJO Phases 6-8), on average the calibrated ECMFS produces enhanced TC activity whose spatial track density distribution agrees well with observations when the initial MJO phase is centered in the Indian Ocean. The exception is the Gulf of Mexico where the observed level of TC activity does not appear to be qualitatively well captured relative to the forecast anomaly magnitude. In terms of maximum forecast probabilities, when the MJO is not centered in the Indian Ocean but has amplitude of at least one standard deviation, maximum forecast probabilities never exceed the low forecast probability range of 10–30%.

One way to compare the uncorrected and calibrated ECMFS and to examine their sensitivity to the MJO is through the reliability diagram. Figure 5.10 shows the reliability diagram of the uncorrected ECMFS, calibrated ECMFS, and calibrated ECMFS but for cases when the initial phase of the MJO is in Phases 1–3 for forecast days 11–32 during the period 2008 to 2011. In terms of the uncorrected ECMFS, forecast probabilities never exceed 30% and agree with observations only 20% of the time. From the frequency diagram inset, rarely does the uncorrected ECMFS produce forecast probabilities greater than 30%, which is in agreement with the findings for the 15-day VarEPS presented in Section 5.3.2. Using the calibrated ECMFS, maximum forecast probabilities reach 40–50%, and these forecast values agree with observations roughly 20–30% of the time. When compared to the uncorrected ECMFS, the calibrated ECMFS generally demonstrate smaller conditional bias especially for forecast probabilities ≤30%. Finally, the reliability of the calibrated ECMFS overall is contrasted with the calibrated ECMFS except for forecasts when the when the MJO is centered in the Indian Ocean at

the time of model initialization. The reliability diagram indicates that for these forecasts, maximum forecast probabilities reach 50–60% and agree with observations 30–40% of the time.

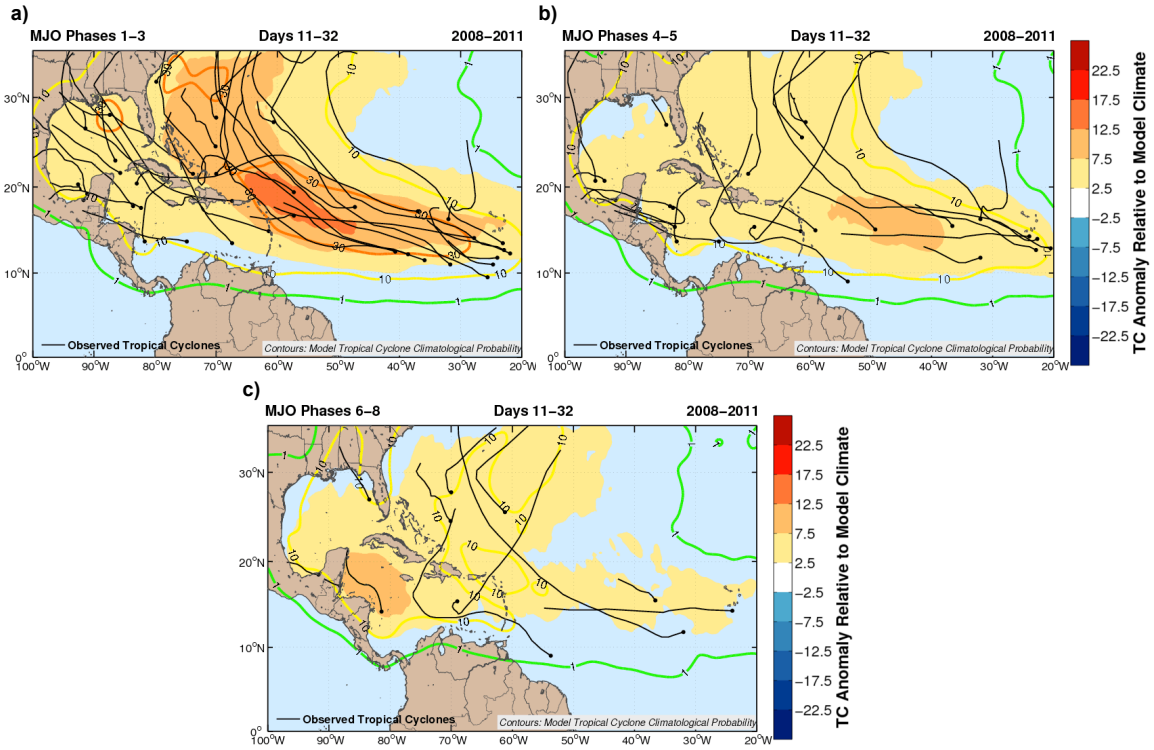

Figure 5.9: Mean track density anomalies (shaded) with bias-corrected probabilities (contoured) based on forecasts when the amplitude of the MJO is greater than one standard deviation for a) MJO Phases 1–3, b) MJO Phases 4–5, and c) MJO Phases 6–8.

To complement the conditional bias view of the uncorrected and calibrated ECMFS on intraseasonal time scales, the relative operating characteristic score for the tropical Atlantic domain defined as 0°N–35°N x 100°W–5°W for the period 2008–2011 is shown in Table 5.4. Since the ROC scores are also stratified by MJO phase and amplitude, the ROC score for climatology is also included as a point-of-reference. When climatology is the benchmark for forecast skill, the uncorrected ECMFS does not produce skillful forecasts. However, the uncorrected ECMFS is skillfull relative to random chance alone as ROC scores are above 0.5 for all forecast dates and MJO phases. Compared to the uncorrected ECMFS, the calibrated ECMFS increases the ROC score on average by 37%. The most significant improvement (40%) is seen with forecasts issued

when the MJO is weak, whereas the smallest improvement (27%) is seen when the active convective phase of the MJO is centered in the Western Hemisphere (Phases 6–8). The calibrated ECMFS also improves upon climatology for the respective forecast period, although the increase in forecast skill is much smaller and averages 4% for all forecast dates. When the relative difference is considered with respect to MJO phases, the largest improvement upon climatology (14%) is seen when the MJO is in Phases 6–8. When the ROC scores are calculated for smaller sub-domains in the tropical Atlantic basin (not shown), the primary regions impacted by this increase in forecast skill is the MDR (37%) followed by the Caribbean Sea (19%). These ROC scores in conjunction with the previous conditional bias results confirm the earlier findings presented in Chapter 4. Namely, the most skillful TC forecasts for the tropical Atlantic on intraseasonal time scales using the ECMFS occur when the active convective phase of the MJO is centered in the Indian Ocean (Phases 1–3) at the time of model initialization. In addition, when the MJO is convectively active but not located in this region, overall the forecast skill of the ECMFS is similar to forecasts issued when the MJO is not active at all.

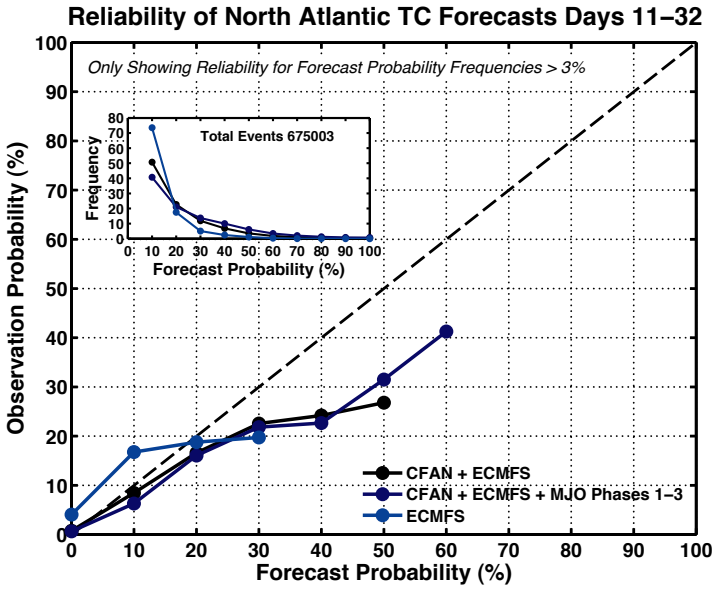

Figure 5.10: Reliability diagram of the ECMWF Monthly Forecast System (ECMFS) for Forecast Days 11-32 in blue, bias-corrected forecast probabilities in black, and bias-corrected forecast probabilities in dark blue for forecasts when the initial phase of the MJO is in Phases 1–3.

Table 5.4: Relative operating characteristic (ROC) score of climatology, uncorrected ECMWF VarEPS, and calibrated ECMWF VarEPS for forecast days 11–32 during the period 2008–2011. The ROC scores have been calculated using a tropical Atlantic domain of 0^{o}N–35^{o}N x 100^{o}W–5^{o}W at a horizontal resolution of 2.5^{o}. The forecasts are evaluated based on initial MJO phase and amplitude, where weak MJO is denoted by amplitude less than one standard deviation. The values in parenthesis list the number of forecasts included in the evaluation. ROC scores greater than 0.5 indicate skill beyond random chance alone.

ROC Score	18-Year Climate	ECMWF VarEPS	Calibrated ECMWF VarEPS
All Dates (88)	0.75	0.57	0.78
MJO 1–3 (24)	0.82	0.61	0.84
MJO 4–5 (15)	0.70	0.52	0.71
MJO 6–8 (9)	0.66	0.59	0.75
Weak MJO (40)	0.73	0.55	0.77

5.4. Conclusions

In this chapter, the forecast components that have been developed thus far for an operational forecast system for North Atlantic tropical cyclones have been reviewed. Overall, the forecast philosophy uses a hybrid technique to generate TC forecasts in which dynamically-based ensemble forecasts from the VarEPS and ECMFS are statistically rendered to remove systematic model bias and to calibrate forecast spread. The forecasts covering the next 15 days are produced twice-daily and utilize a Monte Carlo resampling approach to remove along-track and cross-track biases as a function of forecast lead-time. The most significant improvement in track forecasts is found at short lead-times within 96 hr, with the relative improvement decreasing as forecast lead-time increases. When the Monte Carlo resampling technique is applied to the ECMWF deterministic track forecasts from 2011, the calibrated track forecasts outperform the NHC forecasts on average by 10% beginning at a lead-time of 72 hr and are competitive with the NHC track forecasts from 2010 for lead-times of 48–96 hr. One potential reason

why this approach does not lead to greater improvement at longer lead-times is that the along-track and cross-track error distribution for the 2010–2011 VarEPS forecasts is significantly different than the ECMWF hindcasts. This result may indicate that at longer lead-times the hindcast track projections may not be representative of track biases in the real-time VarEPS.

In addition, the large set of synthetic forecast tracks generated from the resampling approach is used to develop calibrated tropical cyclone track probabilities. The track probabilities are evaluated using the reliability diagram and relative operating characteristic score. When compared to the uncorrected VarEPS, the calibrated VarEPS shows no significant change in conditional bias. ROC scores, however, indicate that the calibrated forecasts provide greater statistical resolution at all forecast lead-times, with the most significant improvement (by 14%) seen during forecast days 4–10. A topic for future work is to evaluate the robustness of these conditional biases and to consider incorporating the forecast probability–observation frequency relationship as an additional post-processing procedure once the track probabilities have been derived from the calibrated VarEPS.

Although the primary focus in the development of the operational forecast system has been in improving probabilistic TC forecasts, a variety of approaches have been evaluated to improve deterministic track forecasts generated from the VarEPS. The current technique utilizes a high-predictability clustering routine to isolate the top five ensembles of the VarEPS whose covariance in forecast tropical cyclone position during the first 72 hr is in best agreement with the ECMWF high-resolution deterministic. The working hypothesis is that this cluster will on average show lower total track errors at extended forecast lead-times than either the mean VarEPS or ECMWF deterministic. When this idea is evaluated for all post-genesis track forecasts during the period 2009–2011, the high-predictable cluster is shown to produce track errors that are on average 25% smaller than the ensemble mean through a lead-time of 144 hr. When compared to

the ECMWF deterministic, the high-predictable cluster does not on average provide a better track forecast. However, this relationship is sensitive to uncertainty in the forecast as indicated by the ensemble spread in the VarEPS. When the condition is made to only consider forecasts with larger than average ensemble spread, the high-predictable cluster is shown to provide track forecasts with the smallest overall error at forecast lead-times beyond 108 hr. In addition, the mean VarEPS becomes more competitive with the ECMWF deterministic but nevertheless continues to provide the forecast with the largest track error at extended lead-times. The take-home point is that the ensemble spread of the VarEPS should be utilized when the goal is to maximize track predictability. For cases when the ensemble spread is lower than normal, the ECMWF deterministic on average will provide a superior track forecast even at extended lead-times. However, when forecast uncertainty is larger than normal during the short-range, a combination of the ECMWF deterministic for short lead-times followed by the high-predictable cluster at extended lead-times will provide the best track forecast on average. A topic of future work is to examine whether intensity forecasts using the high-predictable cluster, ECMWF deterministic, and VarEPS ensemble spread show similar behavior.

To correct the VarEPS maximum intensity forecasts, a quantile-to-quantile mapping technique developed from the ECMWF hindcasts and historical TCs for the period 1992–2010 is used. Calibrated intensity forecasts made during the hurricane seasons of 2010 and 2011 are evaluated using the ranked probability score. RPSs are used to compare the ability of the uncorrected and calibrated VarEPS to forecast correctly the categories of tropical depression, tropical storm, hurricane, and major hurricane at various forecast lead-times. The uncorrected VarEPS shows intensity skill during the first 96 hr, but beyond this point the RPSs are not statistically different from random chance. The calibrated VarEPS shows the most significant improvement relative to the uncorrected VarEPS, on average by 40%, during the first 72 hr. At forecast lead-times

through 192 hr, the RPSs for the calibrated intensity forecasts are improved by 15–20% and show skill beyond random chance alone.

In addition to the twice-daily 15-day outlooks for tropical cyclones in the North Atlantic, a forecast algorithm for the intraseasonal period has also been developed and is used to produce twice-weekly TC forecasts through 32 days. Utilizing the ECMFS, ECMWF hindcasts, observed climatology of tropical cyclones, and the phase and amplitude of the MJO, this scheme uses a similar forecast philosophy as the 15-day outlook. First, the ECMFS and hindcast tracks that originate in the eastern tropical Atlantic are bias-adjusted to their climatological genesis location. Next, the ECMFS intensity forecasts are calibrated using the q-to-q map developed from the ECMWF hindcasts for the 15-day forecast period. Afterwards, track density biases between the ECMWF hindcasts and observed climatology for the overlapping forecast periods are calculated and are used to derive the bias-adjusted forecast track probabilities for the calibrated ECMFS. Finally, track density anomalies for the calibrated ECMFS are derived with respect to the ECMWF hindcast climate.

The calibrated intraseasonal forecasts from the ECMFS during the period 2008–2011 are evaluated using a variety of techniques. First, composite forecast probabilities and track density anomalies are produced for MJO Phases 1–3, 4–5, and 6–8 and confirm the spatial distribution changes that were indicated in the intraseasonal TC predictability study of Chapter 4. Composite track density anomalies for MJO Phases 1–3 are 10–15% above normal for the MDR, stretching northward of the Greater Antilles, and east of the United States, which agrees well with the spatial distribution of observed TC tracks. Maximum calibrated forecast probabilities from the ECMFS also vary with MJO phasing, as peak probabilities reach the range of 30–60% for MJO Phases 1–3, but never exceed the low forecast probability range of 10–30% when the MJO is either convectively active and located elsewhere or inactive. Verification statistics using the reliability diagram indicate that forecast calibration of the ECMFS results in smaller conditional biases

relative to the uncorrected ECMFS and that the most reliable forecasts for the tropical Atlantic basin are produced when the MJO is convectively active and is centered in the Indian Ocean. In this case, gradations of forecast probabilities in the range of 30–60% agree with observations approximately 20–40% of the time.

To evaluate forecast resolution, ROC scores of the uncorrected ECMFS, calibrated ECMFS, and climatology are calculated and reveal that the calibrated ECMFS has the highest ROC scores overall, independent of both MJO phasing and amplitude and shows skill greater than climatology for forecast days 11–32. When interpreted relative to the reliability findings, these results support the delineations of forecast confidence that are provided objectively in the operational forecast system for the intraseasonal period. More importantly, these results provide additional evidence that skillful North Atlantic tropical cyclone forecasts can be provided on intraseasonal time scales. These forecasts are most likely to be deemed useful by end-users who can utilize low (10-30%) and moderate (30-60%) forecast probabilities for tropical cyclones like the energy trading industry or short-term insurance/reinsurance.

CHAPTER 6
INTERANNUAL VARIABILITY OF EASTERLY WAVES AND ITS IMPACT ON NORTH ATLANTIC TROPICAL CYCLONES

6.1. Background and Motivation

A review of previous research on the tropical cyclone–easterly wave relationship on interannual time scales is presented in detail in Section 1.3. Given that 60% of all North Atlantic hurricanes are ultimately triggered by African easterly waves (Agudelo et al. 2011), it is surprising that this relationship has received so little attention on interannual and longer time scales. In recent years, various approaches have been used to derive a climatology of easterly waves, although each technique is not without its deficiencies. In summary, the principal motivation for a new climatology is that previous climatologies by Thorncroft and Hodges (2001) and Hopsch et al. (2007) use a tracking algorithm that detects coherent vortical structures across Africa and the tropical Atlantic that likely includes a mixture of easterly waves and mesoscale convective systems. Agudelo et al. (2011) uses a different approach for easterly wave tracking; however, their climatology is strictly valid for easterly waves that move from Africa into the tropical Atlantic. In addition, all previous climatology studies are based on either the short ERA-15 reanalysis or ERA-40 reanalysis product. Given the marked difference in climatological background structure among reanalyses (see Section 6.3.1), it is instructive to develop separate easterly wave datasets to ensure the climatology of easterly waves is robust.

In summary, this chapter describes the development of a climatology of easterly waves for the East Pacific, for the tropical North Atlantic, and over Africa and presents some initial analysis relevant to the tropical cyclone–easterly wave relationship on

interannual time scales. The historical climatology of easterly waves is determined for several global reanalysis datasets including: the Climate Forecast System-Reanalysis (CFS-R; 1979–2010), ERA-Interim (1980–2010), ERA-40 (1958–2001), and NCEP/NCAR Reanalysis I (1948–2010). The first two datasets represent the current state-of-the-art in terms of global reanalysis but are temporally limited to years encompassing the satellite era. Although the second two datasets may not be as robust, they have longer historical records and may provide a longer climatological perspective of how the easterly wave–tropical cyclone relationship has evolved with time.

This chapter is divided into the following sections. First, in the Data and Methodology of Section 6.2, the easterly wave tracking algorithm is presented which uses curvature vorticity anomalies to derive the trough location of easterly waves. In Sections 6.3.1 and 6.3.2, the various easterly wave datasets constructed from each reanalysis are compared to identify similarities and differences in the easterly wave climatologies. This evaluation is followed by an analysis of the interannual variability in the easterly wave–tropical cyclone relationship based on changes in frequency and easterly wave intensification efficiency with time. Concluding remarks and topics of future work are provided in Section 6.4.

6.2. Data and Methodology

6.2.1. Easterly Wave Tracking Algorithm

Several factors were considered before constructing the easterly wave tracking algorithm used in this analysis. First, the approach by Agudelo et al. (2011) was examined whereby easterly wave tracks are identified using the Hovmöller method based on spatio-temporal filtered 2-6 day westward moving relative vorticity anomalies. This approach produces robust wave trajectories for easterly waves leaving Africa and moving across the tropical Atlantic. However, a number of opportunities for algorithm improvement were identified. First, the original algorithm does not allow for the

continuous tracking of an easterly wave once it has weakened below the tracking threshold but later reintensifes. Second, this method only identifies easterly waves that pass across a particular longitude and therefore is specifically tailored to address easterly waves that move from Africa into the tropical Atlantic. Thus, this approach is not designed to track easterly waves that develop in-situ in the tropical Atlantic or Caribbean or short-lived waves that form and decay over Africa. Third, the use of spatio-temporal filtered variables introduces a phase shift in the Hovmöllers of relative vorticity anomalies especially over Central America, limiting the scheme's ability to identify which easterly waves pass from the Atlantic into the East Pacific. In addition, the use of band-passed filtered data requires long time series making it inappropriate for use in real-time operational wave tracking applications. Finally, the use of relative vorticity anomalies alone makes tracking of easterly waves more difficult through regions where relative vorticity is dominated by background shear vorticity (i.e. the southern Caribbean).

Another easterly wave tracking routine detailed in Berry et al. (2007) was also considered. This scheme uses the advection of streamfunction curvature vorticity to identify the location of ridges or troughs for a particular pressure level and time step. This method has the advantage over Agudelo et al. (2011) by objectively indicating the location of easterly waves regardless of longitude. In addition, the method provides supplemental structural information that the Hovmöller method does not provide, including the meridional wavelength as well as the meridional tilt of the trough axis of the easterly wave, which can indicate whether the easterly wave is growing or decaying. However, the main deficiency of the Berry et al. approach is that it does not provide a way to track the trough locations of easterly waves as a function of time. The Berry tracking method is also ineffective at indicating when an easterly wave has moved into the East Pacific from the Atlantic since the contribution of curvature vorticity from the background climatology is not removed prior to the tracking of easterly waves.

6.2.2. Data Preparation and Wave Identification

In order to combine the benefits of the Agudelo easterly wave tracking method with the Berry trough identification technique, a new tracking scheme is developed and is shown schematically in Figure 6.1. The easterly wave tracking dataset requires the following input variables: the 2D horizontal wind field (**V**), curvature vorticity anomalies (ζ_{cur}), and the advection of curvature vorticity anomalies ($-\boldsymbol{V}\cdot\boldsymbol{\nabla}_h\zeta_{cur}$) all at a particular pressure level. Since curvature relative vorticity, especially in the lower troposphere, is often influenced by local topography across Central America and over Africa, curvature vorticity anomalies at a pressure level are calculated by removing the 6-hr long-term average. The 6-hr climatology varies with the reanalysis product used for easterly wave tracking and is defined as the temporal extent of the reanalysis (i.e. CFS-R: 1979–2010, ERA-40: 1958–2001, ERA-Interim: 1980–2010, NCEP/NCAR Reanalysis I: 1948–2010). The key variables from these reanalysis datasets are decimated onto a regular latitude-longitude geographic coordinate system with a horizontal resolution of 2.5^o x 2.5^o. Next, each variable is smoothed twice using a 9-point local smoother, which helps to reduce vorticity signals from mesoscale convective systems while retaining the synoptic-scale easterly wave structure. Finally, easterly wave trough locations are defined to occur in regions where $U < 2.5\ m\ s^{-1}$, $\zeta_{cur} > 0.25 \times 10^{-5}\ s^{-1}$ and $-\boldsymbol{V}\cdot\boldsymbol{\nabla}_h\zeta_{cur} = 0$.

6.2.3. Wave Evaluation and Merging

After identifying the trough locations of easterly waves, the next step is to ensure that the delineated trough axes are unique waves. Each trough axis is compared to neighboring trough axes to determine whether the trough axis belongs to the same wave. This wave merging step uses the median latitude and longitude of the wave axis along with recursion to isolate the locus of points surrounding this location that exceed the wave trough minimum curvature vorticity threshold ($\zeta_{cur\ T} > 0.5 \times 10^{-5}\ s^{-1}$). However, if the latitudinal or longitudinal extent of the isolated region is in excess of 25^o, then the ζ_{cur}

$_T$ is increased by 50% until the enclosed region is smaller than 25^o. Next, a polygon is constructed using the outer locus of selected points. Then, the median latitude and longitude coordinates for each trough axis is compared to a particular trough axis polygon to determine if these coordinates lie inside or outside of the polygon. Any trough axis that is located within the polygon enclosing the wave axis is defined as belonging to the same easterly wave and is merged. During the merging process, a new average latitude and longitude is defined which is the average of all median latitude and longitude trough locations as well as the latitude and longitude extent of the trough axes. This process results in the creation of a new subset of merged trough axes. Next, a simple check is applied to ensure that no duplicate trough axes are found within the merged dataset. In this case, any trough axes that are located within 300 n mi are assumed to belong to the same easterly wave and are accordingly merged.

6.2.4. Wave Tracking and Post-Processing

After the wave merging process is complete, the next step is to compare the current set of waves with any easterly waves that are present at the previous time step for time $t > t_0$. If wave tracking is at the initial time step, then the current set of wave troughs becomes the first set of finalized easterly waves. For each easterly wave, identification information that is stored includes ξ_{max}, the wave's centroid latitude and location, ξ_{avg} of the trough region of the easterly wave, and the minimum and maximum latitudinal and longitudinal wave extent. Next, a 6-hr and 12-hr position estimate is calculated based on the assumption that the easterly waves movement is based on advection from the environmental flow. These position estimates are critical in accurately identifying the future location of an easterly wave.

For time $t > t_0$, the wave tracking methodology proceeds first with the identification of wave trough locations using the methodology defined in Section 6.2.2. Next, a subset of unique wave troughs is calculated using the wave merging process

defined in Section 6.2.3. Afterwards, each 6-hr wave position estimate from the previous time step is compared to the set of unique wave troughs from the current time step. In this case, the wave trough whose great-circle distance between its current location and the estimated 6-hr position for a previous easterly wave is minimized and is less than or equal to 600 n mi is defined as the current position of an easterly wave. For these waves, identification information is stored and new 6-hr and 12-hr position estimates are made for likely wave locations at future time steps. For easterly waves from 12 hr earlier that do not have a matching wave location at the current time step, the wave tracking methodology uses the great-circle distance between the estimated 12-hr wave position and the current set of wave troughs. A wave match occurs for the current wave trough so long as the distance separation between the estimated position of the observed location is less than or equal to 900 n mi. Using current wave troughs that have not been matched to earlier easterly waves using 6-hr or 12-hr position estimates, these systems are classified as new wave troughs and identification information is stored in the easterly wave dataset along with position estimates for 6 hr and 12 hr after the current time step. Once wave tracking has been completed, the easterly wave dataset undergoes additional post-processing by requiring that the easterly wave's lifetime, Δt_{life}, is at least one day and that the easterly wave has tracked westward over its lifetime, Δlon_{life}, by at least 2.5^{o}. This post-processing step is considered a weak constraint on the easterly wave climatology. However, for the results that are shown in Section 6.3, the easterly waves must have a lifetime of at least four days.

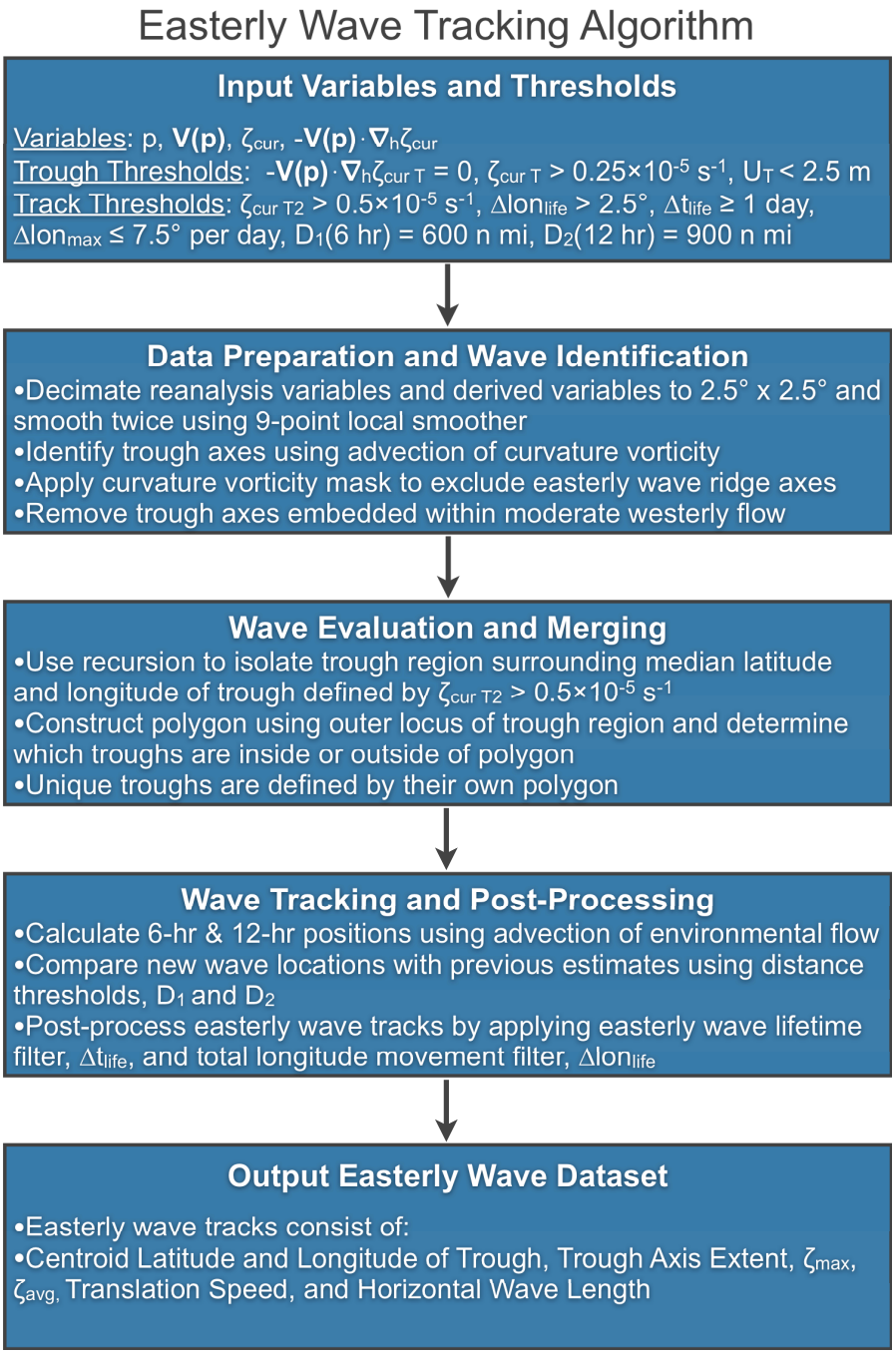

Figure 6.1: Schematic of the easterly wave tracking algorithm that uses zonal and meridional winds, curvature vorticity anomalies, and the advection of curvature vorticity anomalies for a particular pressure level.

6.3. Results

6.3.1. Comparison of Reanalysis Datasets

Figure 6.2 shows the climatological mean state at 700 hPa over the tropical Atlantic and eastern Pacific in terms of relative vorticity and total winds for the ERA-Interim, ERA-40, CFS-R, and NCEP/NCAR reanalyses during the period July to September in the overlapping period of 1980–2001.

All reanalyses show varying degrees of strength with the mid-tropospheric African Easterly Jet (AEJ), which is reflected in the north-south dipole of relative vorticity that extends from western Africa into the eastern tropical Atlantic. Relative to the ERA-40 reproduction of the AEJ, the CFS-R depicts a stronger dipole which is also reflected in a stronger and greater westward expansion of the mid-tropospheric easterly jet into the tropical Atlantic. In addition, the AEJ over Africa in the CFS-R is displaced a few degrees southward relative to its position in ERA-Interim. These differences in the climatological mean state have important implications on how easterly waves are reproduced within each reanalysis. Leroux and Hall (2009) show that for a given source of convective heating, the structure and configuration of the AEJ has a profound impact on the easterly wave response with stronger easterly waves characterized by a strong AEJ, strong vertical wind shear, or an elongated potential vorticity reversal from Africa through the eastern tropical Atlantic.

One may speculate that reanalyses with a greater westward expansion of the AEJ across the tropical Atlantic will result in more persistent barotropic instability, providing African easterly waves greater opportunity to convert zonal available potential energy to eddy kinetic energy. In addition, the southward shift of the AEJ in the CFS-R should lead to a greater frequency of southern-tracked African easterly waves with larger amplitudes relative to the ERA-Interim reanalysis. The implications of these differences

in the mean state are considered further in the assessment of the seasonal climatology of easterly waves in Section 6.3.2.

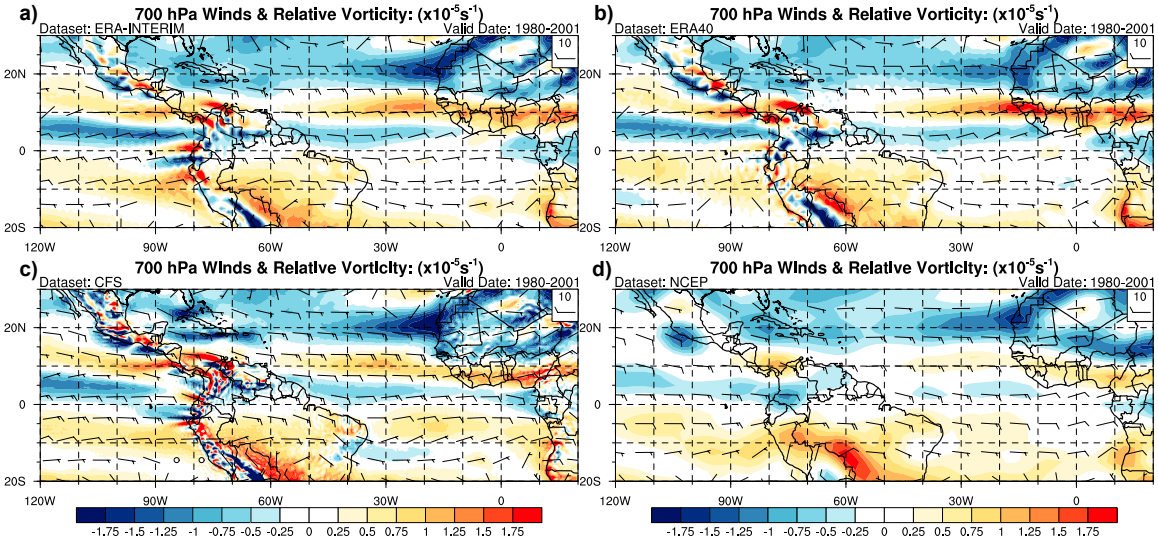

Figure 6.2: Average 6-hr 700 hPa winds and relative vorticity for the period July–September from 1980 to 2001 for a) ERA-Interim, b) ERA-40, c) CFS-R, and d) NCEP/NCAR Reanalysis.

Figure 6.3 depicts the monthly average 6-hr variance in relative vorticity at 700 hPa for the period July to September from 1980 to 2001 after translating each reanalysis to a regular 2.5^{o} x 2.5^{o} latitude-longitude grid. At first glance, the ERA-Interim, ERA-40, and NCEP/NCAR reanalyses are quantitatively similar in the spatial distribution of maximum and minimum relative vorticity variance across the East Pacific and tropical Atlantic, whereas the CFS-R shows larger variability in the regions of highest overall variance. In particular, all four reanalyses show that northwestern Africa has greater variability at 700 hPa than southwestern Africa. However, ERA-40 shows greater variability in relative vorticity across the African Sahel than either NCEP/NCAR or ERA-Interim and is in better agreement with the spatial distribution found in the CFS-R. Across the East Pacific, the differences among the reanalyses are similar to that found across the tropical Atlantic: the CFS-R depicts the greatest variance in mid-tropospheric relative vorticity; ERA-Interim and ERA-40 are similar quantitatively and in overall distribution, whereas NCEP/NCAR systematically produces reduced relative vorticity

variability in this region. When these differences among reanalyses are considered with the comparison of the mean base state, these discrepancies should yield unique depictions of the frequency or distribution of easterly waves with each reanalysis. In particular, one would expect the East Pacific easterly wave tracks to be similar between the ERA-40 and ERA-Interim reanalyses given the similar distribution of 700 hPa relative vorticity variance. Also, the CFS-R is likely to produce either the largest number of easterly waves or the easterly waves with the greatest amplitude in both the East Pacific and tropical Atlantic.

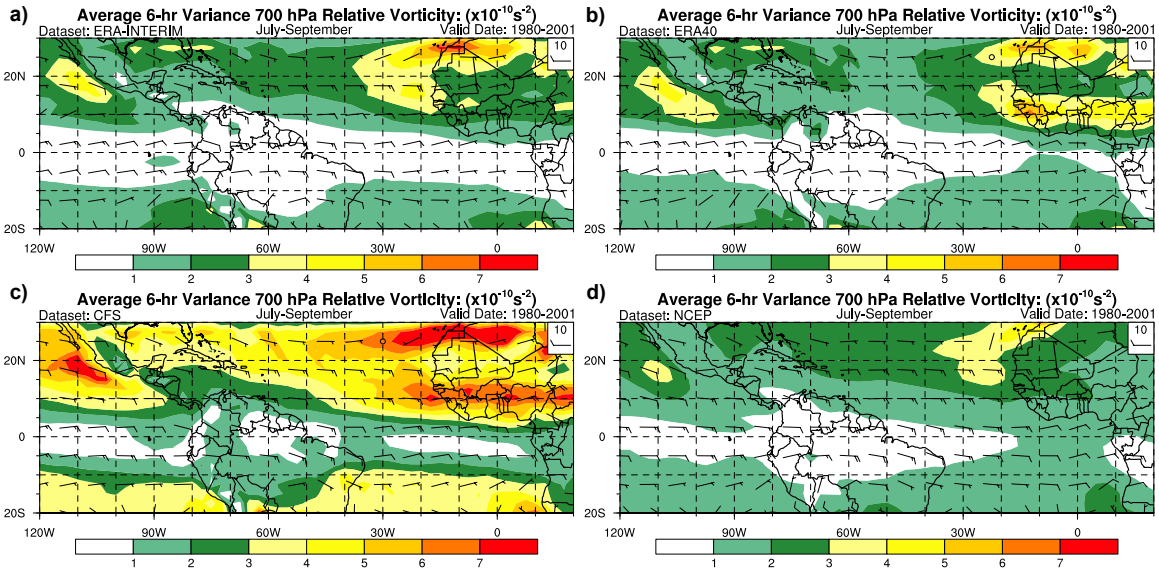

Figure 6.3: Similar to Figure 6.2, except for 6-hr variance for 700 hPa winds and relative vorticity for the period July–September from 1980 to 2001 for a) ERA-Interim, b) ERA-40, c) CFS-R, and d) NCEP/NCAR Reanalysis.

In the previous comparisons, the differences among the reanalyses have been considered in terms of how mid-tropospheric relative vorticity is reproduced. Figure 6.4 summarizes the differences between the ERA-Interim and CFS-R using the correlation coefficient of 6-hr relative vorticity at 600, 700, and 850 hPa after interpolating each reanalysis to a regular 2.5^{o} x 2.5^{o} grid. The greatest similarity between the two reanalyses in terms of low to mid-tropospheric relative vorticity is found over the western tropical Atlantic, across the Greater Antilles, and in the Gulf of Mexico, where about

50% of the total variance may be explained by each respective reanalysis. However, the weakest agreement between the reanalyses is located in the ITCZ, over much of equatorial Africa, and the tropical South Atlantic. These regions share the common characteristic of observation paucity, as the ITCZ is very difficult to sample using remote sensing techniques due to persistent, deep convection, and equatorial Africa lacks rawinsonde observations. The lack of agreement in these regions suggests that the atmospheric model used to produce each reanalysis will in a sense be generating its own reality in terms of easterly wave structure, amplitude, and lifetime due to rather weak constraints from the limited spatial and temporal distribution of observations.

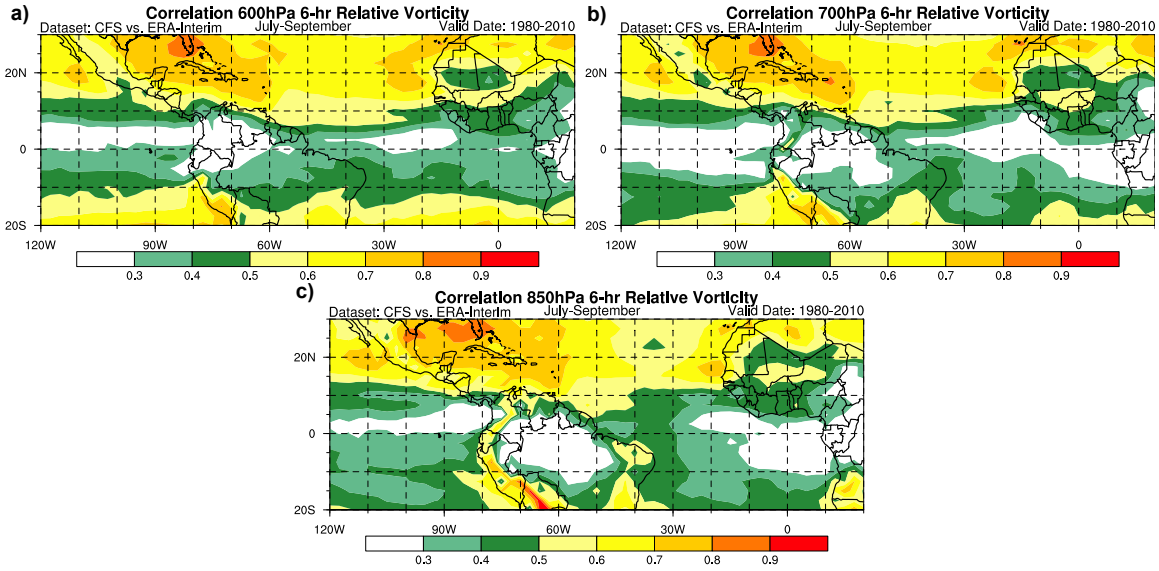

Figure 6.4: Correlation coefficient of 6-hr relative vorticity between the ERA-Interim and CFS-R for the period 1980–2010 after interpolating both reanalyses to a 2.5° x 2.5° regular latitude–longitude grid.

6.3.2. Seasonal Climatology of Easterly Waves

Figure 6.5 shows the seasonal variability of 700 hPa easterly waves that moved through the MDR and the Caribbean Sea (10°N–20°N x 20°W–80°W) and had a lifetime of at least four days for each atmospheric reanalysis dataset for the overlapping period of 1980–2001. Results from the easterly wave tracking algorithm reveal that easterly waves move through the tropical Atlantic during all months of the year with increased frequency

during the months May to October and with peak easterly wave activity found during August. As discussed in Agudelo et al. (2011), these structures found during the boreal winter are not convectively coupled and therefore would not be detected if easterly wave tracking were based on only outgoing longwave radiation. In addition, as indicated by the spread in the interpercentile distributions, the interannual variability in the total number of easterly waves relative to the mean of the reanalysis periods is quite pronounced, raising the possibility that tropical and extratropical climate variability may be modulating the seasonal frequency of easterly waves.

Generally there is fairly good agreement among the ERA-Interim, ERA-40, and CFS-R datasets for the frequency of waves by month especially during the period May to October. However, the frequency of easterly waves in the NCEP/NCAR reanalysis is 1.5 to 2 times as large as the other reanalysis products during boreal winter. Likewise, during the boreal summer, the NCEP/NCAR reanalysis generates too many waves relative to the other reanalyses and exhibits the greatest intraseasonal variability with a 50% interval of 20 to 47 waves for August relative to about 20 to 40 waves in the other reanalysis datasets. Relative to the period May to October of 1980–2001, the average percentage of easterly waves generated in the tropical Atlantic versus over Africa is 46% for the CFS-R, 45% for the ERA-Interim, 44% for the ERA-40, and 49% for the NCEP/NCAR reanalyses.

Figure 6.6 shows the monthly climatology of easterly waves for the East Pacific. The frequency of easterly waves in the East Pacific is about 2 to 3 times as large as the number of easterly waves in the tropical Atlantic, depending on the reanalysis and month of the year considered. Given the significantly larger number of easterly waves in the East Pacific relative to the tropical Atlantic, this difference is statistical evidence that a large proportion of easterly waves in the East Pacific are actually generated in-situ rather than being advected into the East Pacific from the tropical Atlantic. In fact, relative to the period May to October of 1980–2001, the average percentage of easterly waves

generated in-situ versus being advected into the region is 63% for the CFS-R, 57% for the ERA-Interim, 58% for the ERA-40, and 41% for the NCEP/NCAR reanalyses. These results are in agreement with recent findings by Serra et al. (2008) and Toma and Webster (I,II 2010) and challenge the current paradigm that most East Pacific easterly waves originate from the tropical Atlantic. In addition, although all reanalysis products show proper phasing of easterly wave frequency with most activity during the boreal summer, there are disagreements regarding which month exhibits peak easterly wave activity. For the ERA-Interim and ERA-40 reanalyses, the most frequent easterly waves are seen during July, whereas CFS-R and NCEP/NCAR suggests August has the highest frequency of wave activity.

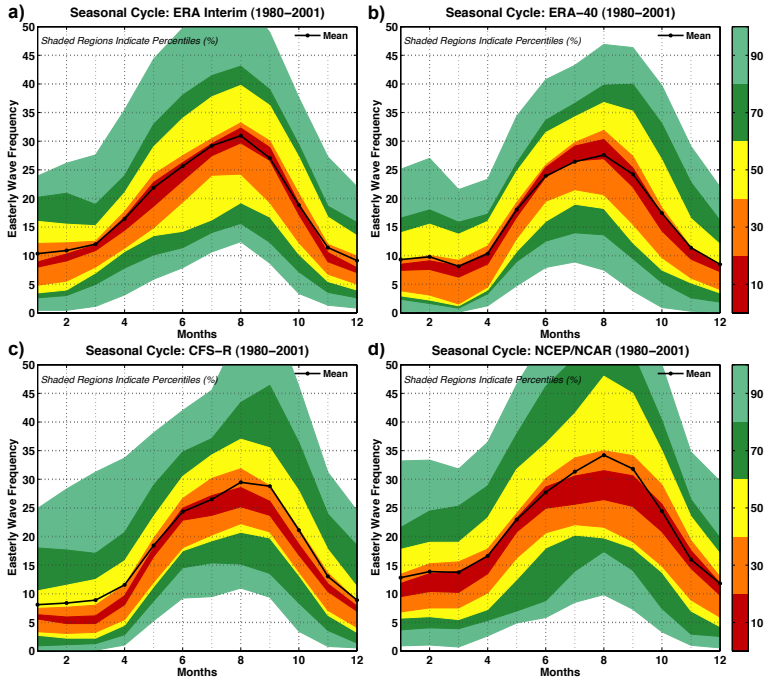

Figure 6.5: Interpercentile plumes of the seasonal frequency of 700 hPa easterly waves in the Atlantic for the a) ERA-Interim, b) ERA-40, c) CFS-R, and d) NCEP/NCAR reanalysis datasets for the overlapping period 1980–2001. Time series have been smoothed with a 1-2-1 filter.

Figure 6.7 shows the track density for 700 hPa easterly waves that originate over Africa and had a lifetime of at least four days for the overlapping period of 1980–2001. Two preferred regions of easterly wave track activity are seen in all reanalysis products:

the primary track density region is oriented southeast to west-northwest across the African Sahel and extending into the tropical Atlantic in association with the mid-tropospheric tropical easterly jet with maximum track density concentrated near Dakar (15^{o}N x 18^{o}W). Both the ERA-40 and NCEP/NCAR reanalyses indicate a higher frequency of easterly wave tracks originating to the north-northwest of Lake Victoria, which is only weakly evident in the ERA-Interim and CFS-R products. In addition, the African easterly wave track density from the NCEP/NCAR reanalysis indicates a more longitudinally oriented track density region through the tropical Atlantic, with significantly higher frequency of African easterly waves moving through the tropical Atlantic and into the East Pacific. The second preferred African easterly wave track region is located in the southern hemisphere in the region 5^{o}S–15^{o}S x 40^{o}W–20^{o}E. The spatial distribution of track density among the ERA-Interim, ERA-40, and CFS-R are in the best agreement in this region with the NCEP/NCAR reanalysis, indicating much greater longitudinal extent both through South America and farther east into eastern Africa.

Figure 6.8 shows 700 hPa easterly wave track density for all easterly waves that moved through the tropical Atlantic and had a lifetime of at least four days. Given the greater track density of Atlantic easterly waves across the region 10^{o}N–25^{o}N x 80^{o}W–20^{o}W relative to the wave density of African easterly waves for this region, this finding indicates that roughly half are generated within the tropical Atlantic, most likely within the ITCZ or via extratropical trough interactions in the region. Among the reanalyses considered here, the ERA-Interim and CFS-R show the greatest spatial similarity in terms of Atlantic easterly wave track density, with the NCEP/NCAR reanalysis, indicating a significantly larger fraction of easterly waves in this region eventually moving into the East Pacific.

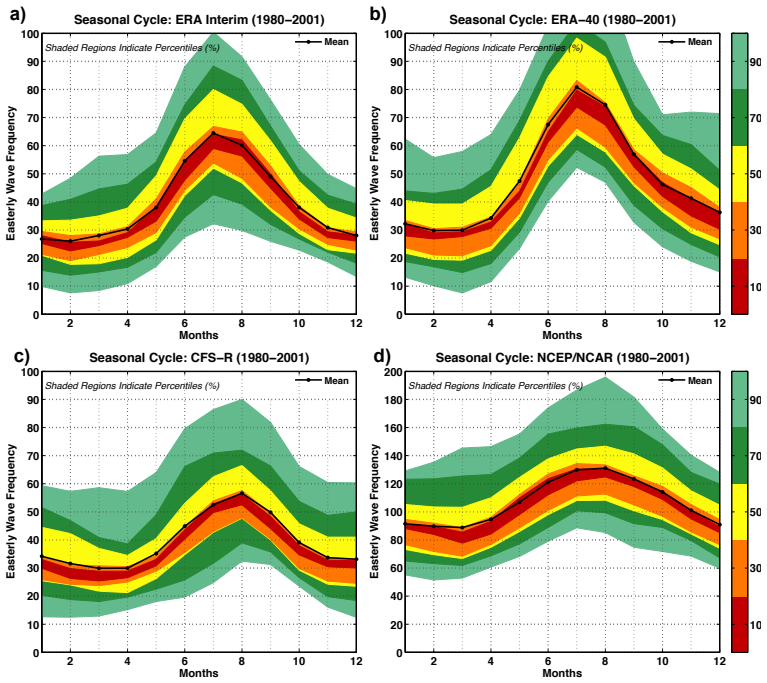

Figure 6.6: Interpercentile plumes of the seasonal frequency of 700 hPa easterly waves in the East Pacific from the a) ERA-Interim, b) ERA-40, c) CFS-R, and d) NCEP/NCAR reanalysis datasets for the overlapping period 1980–2001. Time series have been smoothed with a 1-2-1 filter.

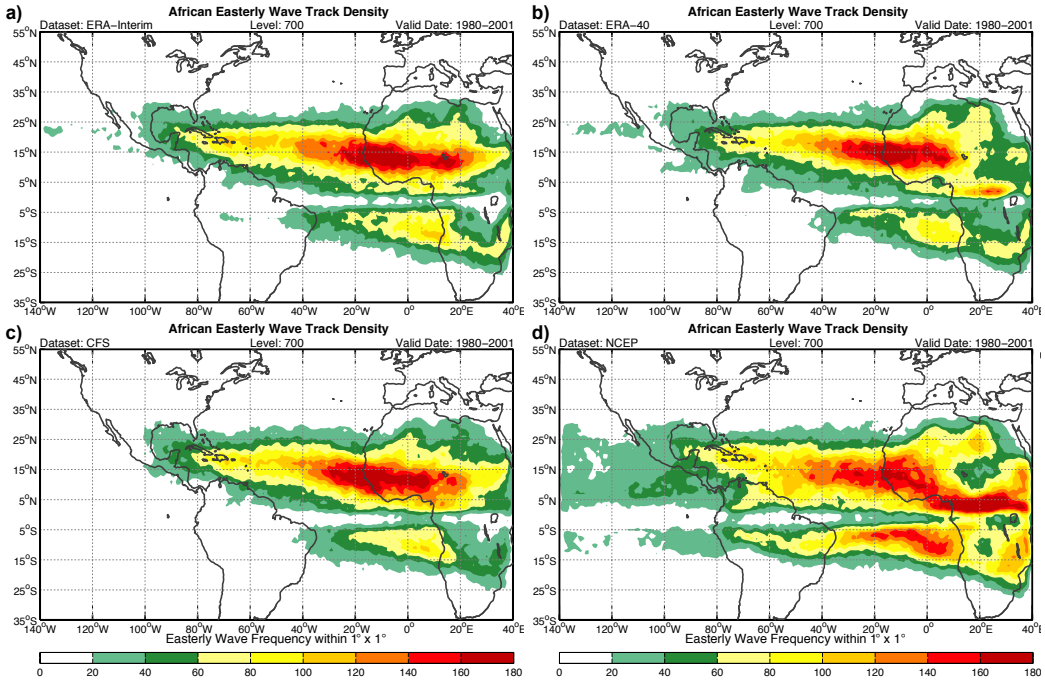

Figure 6.7: African easterly wave track density at 700 hPa for May to October during the overlapping period 1980–2001 for the a) ERA-Interim, b) ERA-40, c) CFS-R, and d) NCEP/NCAR reanalyses.

Besides examining the spatial distribution of easterly wave track density, Table 6.1 summarizes some basic characteristics of easterly waves for each reanalysis. First, the average curvature vorticity within the trough region of tracked easterly waves averages between 9 x 10^{-6} s^{-1} and 1 x 10^{-5} s^{-1} with the maximum curvature vorticity averaging anywhere from 50 to 90% larger than the average easterly wave curvature vorticity. The CFS-R produces the easterly waves with the largest curvature vorticity amplitude, which is consistent with expectations based on the mean and variance base state as well as the higher horizontal resolution of the model's integration. The average phase speed of Atlantic easterly waves is between 13 m s^{-1} and 16 m s^{-1}, which is primarily due to the zonal phase speed of the easterly waves which average between 9-10 m s^{-1}. Finally, the average meridional wavelength of the tracked easterly waves is between 1400 and 1600 km, which confirms that the tracked vortical structures have synoptic spatial scale.

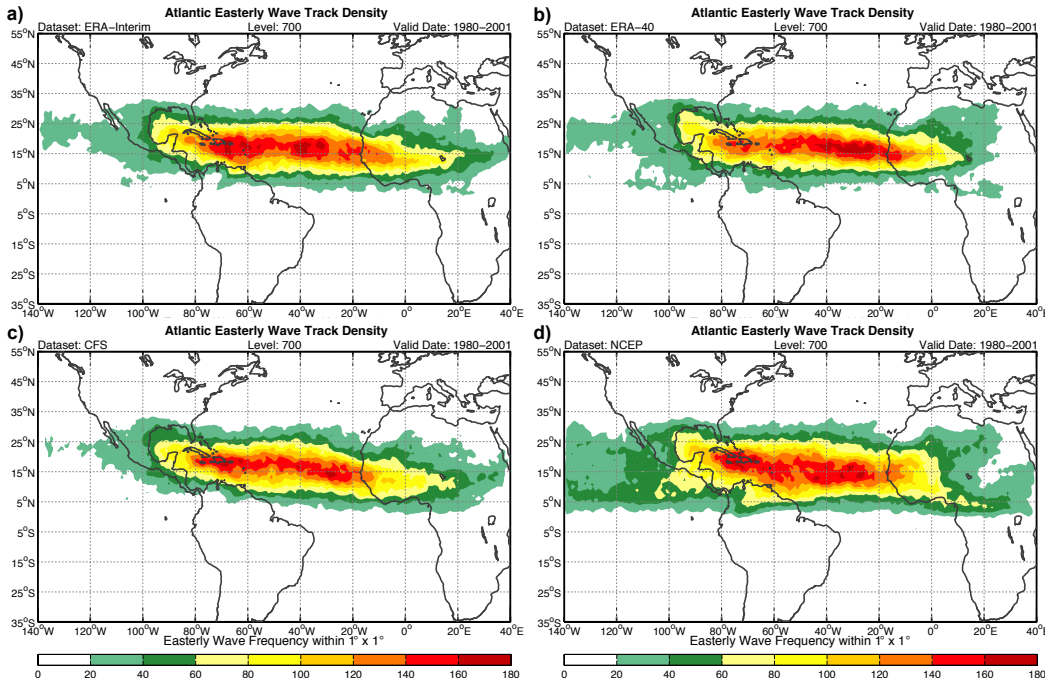

Figure 6.8: Atlantic easterly wave track density at 700 hPa for May to October during the overlapping period 1980–2001 for the a) ERA-Interim, b) ERA-40, c) CFS-R, and d) NCEP/NCAR reanalyses.

Table 6.1: Structural characteristics of Atlantic easterly waves during the period June to October 1980–2001 for the ERA-Interim, ERA-40, CFS-R, and NCEP/NCAR reanalyses.

1980–2001	ERA-Interim	ERA-40	CFS-R	NCEP/NCAR
$\overline{\xi_{cur}}$	9.2 x 10^{-6} s^{-1}	8.9 x 10^{-6} s^{-1}	1.0 x 10^{-5} s^{-1}	1.0 x 10^{-5} s^{-1}
$\overline{Max(\xi_{cur})}$	1.5 x 10^{-5} s^{-1}	1.5 x 10^{-5} s^{-1}	1.9 x 10^{-5} s^{-1}	1.7 x 10^{-5} s^{-1}
$\overline{C_T}$	13.4 m s^{-1}	14.0 m s^{-1}	14.5 m s^{-1}	15.9 m s^{-1}
$\overline{C_Z}$	9.4 m s^{-1}	9.7 m s^{-1}	10.1 m s^{-1}	10.7 m s^{-1}
$\overline{C_M}$	7.7 m s^{-1}	8.2 m s^{-1}	8.4 m s^{-1}	9.6 m s^{-1}
$\overline{\lambda_M}$	1503 km	1463 km	1573 km	1639 km
$\overline{(x, y)}$	18.8N x 52.5W	19.0N x 52.2W	18.5N x 52.6W	18.5N x 53.2W

Figure 6.9 shows the magnitude of easterly wave track density using the overlapping period of 1980–2001 for the peak months of easterly wave activity in the East Pacific, May to October. In this case, the track density is calculated using easterly waves that move through the East Pacific domain and have a lifetime of at least four days. In terms of spatial similarities, all reanalyses indicate two primary wave track regions both north and south of the equator with a latitudinal displacement of 5^{o}, and similar latitudinal extent for the northern (25^{o}) and southern (10^{o}) regions. The greatest similarity among the reanalyses is found with ERA-Interim (Figure 6.9a) and CFS-R (Figure 6.9c). Both reanalyses show the maximum eastward longitudinal extent of easterly wave tracks moving into the East Pacific from the tropical Atlantic is the Lesser Antilles. Both the ERA-40 and NCEP/NCAR reanalyses have a higher frequency of easterly waves originating over Africa and ultimately propagating into the East Pacific. In addition, the NCEP/NCAR reanalysis exhibits the highest track density, which is a factor of two larger than the other reanalyses.

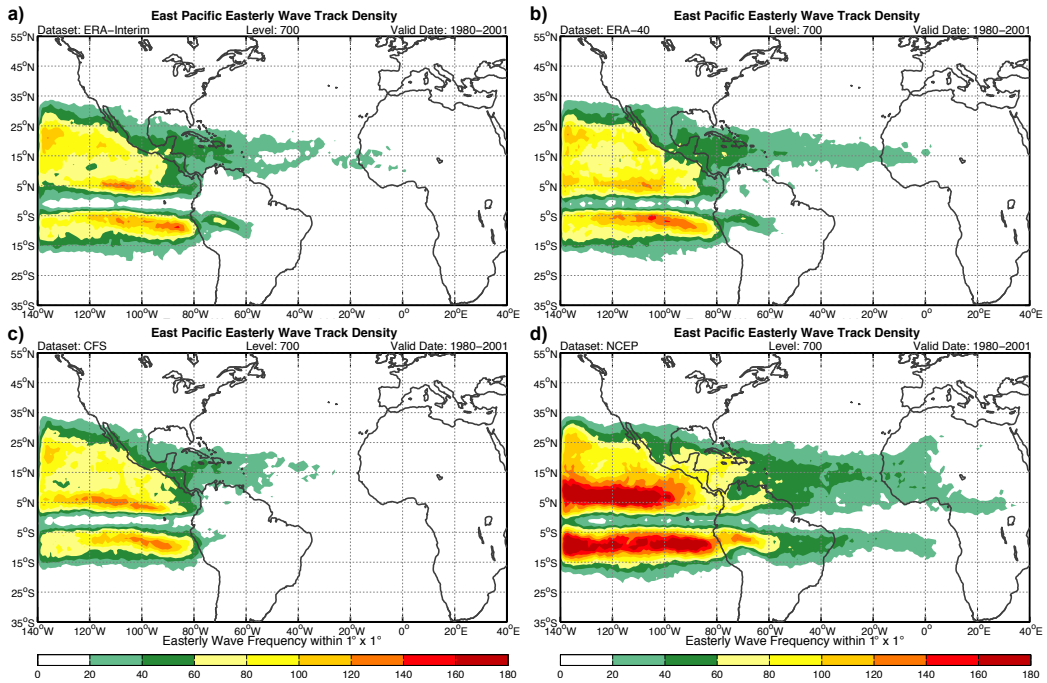

Figure 6.9: East Pacific easterly wave track density at 700 hPa for May to October during the overlapping period 1980–2001 for the a) ERA-Interim, b) ERA-40, c) CFS-R, and d) NCEP/NCAR reanalyses.

Table 6.2 summarizes some basic characteristics of easterly waves that moved through the East Pacific during the period June to October 1980–2001 for the four reanalyses. The average curvature vorticity of the trough region of East Pacific easterly waves is about 10–20% weaker than Atlantic easterly waves and averages 8.6 x 10^{-6} s^{-1} with average maximum curvature vorticity of 1.5 x 10^{-5} s^{-1}. The average phase speed of East Pacific easterly waves is 13.5 m s^{-1}, with an average zonal (meridional) phase speed of 9.5 m s^{-1} (7.6 m s^{-1}). The average meridional wavelength is 1210 km, which is approximately 30% smaller than the average meridional wavelength of Atlantic easterly waves.

Table 6.2: Structural characteristics of East Pacific easterly waves during the period June to October 1980–2001 for the ERA-Interim, ERA-40, CFS-R, and NCEP/NCAR reanalyses.

1980–2001	ERA-Interim	ERA-40	CFS-R	NCEP/NCAR
$\overline{\xi_{cur}}$	8.0 x 10^{-6} s^{-1}	8.1 x 10^{-6} s^{-1}	9.8 x 10^{-6} s^{-1}	8.3 x 10^{-6} s^{-1}
$\overline{Max(\xi_{cur})}$	1.3 x 10^{-5} s^{-1}	1.3 x 10^{-5} s^{-1}	1.8 x 10^{-5} s^{-1}	1.4 x 10^{-5} s^{-1}
$\overline{C_T}$	12.3 m s^{-1}	12.9 m s^{-1}	13.5 m s^{-1}	15.3 m s^{-1}
$\overline{C_Z}$	8.7 m s^{-1}	9.0 m s^{-1}	9.6 m s^{-1}	10.5 m s^{-1}
$\overline{C_M}$	6.8 m s^{-1}	7.2 m s^{-1}	7.4 m s^{-1}	8.9 m s^{-1}
$\overline{\lambda_M}$	1128 km	1148 km	1233 km	1329 km
$\overline{(x, y)}$	7.1N x 113.8W	6.3N x 114.2W	6.9N x 113.4W	5.2N x 114.8W

6.3.3. Easterly Wave–Tropical Cyclone Relationship

One of the motivating reasons to construct the easterly wave climatology is to assess how variability in easterly wave frequency or intensity on interannual time scales is related to observed changes in Atlantic TC activity. Table 6.3 shows the correlation coefficient between the number of Atlantic easterly waves in each reanalysis and tropical cyclone frequency as a function of easterly wave lifetime. With the exception of the CFS-R, all reanalysis datasets show a moderate relationship (correlations between 0.4-0.6) between easterly wave frequency and tropical cyclone activity, which is statistically significant at the 95% confidence level using a boot-strap resampling test. This relationship, however, is sensitive to easterly wave lifetime. For instance, the strongest relationship between ERA-Interim easterly wave frequency and tropical cyclones on annual time scales is for easterly waves with a lifetime of at least eight days, whereas ERA-40 and NCEP/NCAR reanalyses exhibit the strongest relationship for easterly wave lifetimes of at least four and six days, respectively.

The sensitivity in the easterly wave–tropical cyclone frequency relationship to lifetime may be reflective of the inherent reanalysis differences in preferred easterly wave genesis regions over Africa or within the ITCZ, which is also reflected in differences in easterly track density (Figure 6.7 and Figure 6.8) along with differences in the background climate state (Figure 6.2). In addition, the relationship between easterly waves and tropical cyclones in the ERA-40 reanalysis is in contrast to Hopsch et al. (2007) in which their vorticity tracking results showed very weak covariability with Atlantic TCs. However, these new findings confirm the relationship they identified linking 2-6 day filtered meridional winds over the eastern tropical Atlantic, which is a considered a proxy for easterly wave activity and Atlantic TCs.

The easterly wave–tropical cyclone frequency relationship also exhibits spectral covariability. Figure 6.10a shows the amplitude of the cospectrum between Atlantic easterly waves and tropical cyclones using an average of the easterly wave datasets from the four reanalyses for the period 1948–2010 and a 20-year Hanning window for tapering. At time scales of four to five years and at low frequencies of 20 years, the relationship between easterly waves and tropical cyclone frequency is strongest. The coherence squared statistic indicates that the two times series are most coherent and statistically significant beyond the 95% confidence level at time scales of two and four years (Figure 6.10b). The relationship is not statistically significant at 20-year periods, likely due to the limited number of degrees of freedom from the rather short easterly wave data record. Figure 6.11 is a time series of standardized anomalies of the number of easterly waves for each reanalysis and the number of Atlantic TCs after using a low-pass butterworth filter with three weights and a cut-off period of four years. From Table 6.3, the correlation coefficient using the filtered and detrended-filtered datasets is larger than using the annual time series, with the highest correlation coefficient found with the ERA-Interim reanalysis at 0.73. Given that the correlation coefficient for ERA-Interim is about one-third lower using the detrended-filtered time series, this indicates that a large

percentage of the shared variance between easterly waves and TCs is tied to the increasing trend that has been observed in the Atlantic basin since at least 1980 in both easterly wave and tropical cyclone frequency.

Another interesting aspect of the temporal variability between tropical cyclone frequency and easterly waves is the phasing of the lead-lag relationship as a function of time. For instance, from 1963 to 1970 and from 1985 to 1995, the time series of the number of easterly waves leads tropical cyclone frequency by about two years. Then, from 1971 to 1977 and 1996 to 2003, the time series of tropical cyclone frequency and easterly wave activity are in phase with one another. The relationship changes phase with the number of tropical cyclones leading the frequency of easterly waves by about one year for the periods 1978-1984 and from 2004 to present. Since a majority of easterly waves in the tropical Atlantic are generated over Africa, the difference in phase relationship may suggest low-frequency internal variability over Africa is leading to constructive or destructive interference in the easterly wave–tropical cyclone relationship downstream in the tropical Atlantic. However, the physical explanation for this difference is beyond the scope of the current analysis.

Table 6.3: Correlation coefficients between the number of easterly waves as a function of lifetime in the tropical Atlantic basin and tropical Atlantic TC frequency for June to October for each reanalyses respective period of record. Values in bold are statistically significant at the 95% confidence level using a boot-strap resampling technique. The first column, Raw, denotes the annual correlation coefficient. The second column, Fil, lists the filtered correlation coefficient using a butterworth filter with three weights and a cut-off period of four years. The third column, DeFil, lists the correlation coefficient after detrending the filtered time series.

Correlation Coefficient	Easterly Wave Lifetime ≥ 4 Days			Easterly Wave Lifetime ≥ 6 Days			Easterly Wave Lifetime ≥ 8 Days		
	Raw	Fil	DeFil	Raw	Fil	DeFil	Raw	Fil	DeFil
ERA-Interim	0.06	**0.42**	-0.06	**0.45**	**0.60**	0.06	**0.53**	**0.73**	**0.49**
CFS-R	0.03	-0.16	**-0.43**	0.10	-0.01	-0.24	0.21	0.24	0.10
ERA-40	**0.40**	**0.66**	**0.67**	0.26	**0.48**	**0.50**	0.10	**0.33**	**0.38**
NCEP/NCAR	**0.28**	**0.38**	**0.34**	**0.44**	**0.58**	**0.56**	**0.34**	**0.56**	**0.54**

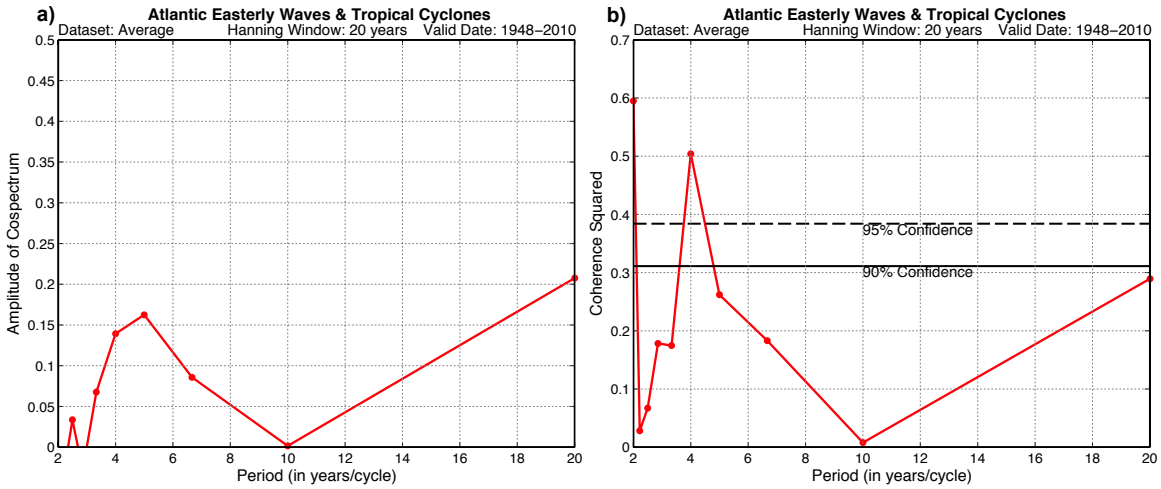

Figure 6.10: a) Cospectrum as a function of period between Atlantic easterly waves and the number of tropical cyclones using a 20-year Hanning window. The easterly wave dataset is an average of the four easterly wave reanalyses. b) Similar to Figure 6.10a, except for coherence squared. The 90% and 95% confidence levels are determined using the F-statistic.

The interannual variability in easterly wave frequency is one possible dynamical mechanism to explain why the number of tropical cyclones in the Atlantic basin has varied historically. However, this hypothesis assumes that the proportion of easterly waves inducing tropical cyclones has remained constant with time. Therefore, it is important to also consider the variability in easterly wave intensification efficiency, which as defined by Agudelo et al. (2011) is the percentage of all easterly waves in a given season that are responsible for tropical cyclone formation. Agudelo et al. (2011) found that there has been no overall trend in the intensification efficiency at least for the period 1980–2001 in the ERA-40 reanalysis. However, their tracking approach only targeted AEWs and used a short period of record based on one reanalysis dataset; alternatively, here we consider the interannual variability in easterly wave intensification efficiency using the easterly wave climatologies for the CFS-R, ERA-Interim, ERA-40, and NCEP/NCAR reanalyses. Figure 6.12 shows the intensification efficiency of Atlantic easterly waves for the period June to October for easterly waves with a lifetime of at least four days. Among the reanalyses, there is very good agreement in the annual variability in easterly wave intensification efficiency as the average inter-reanalysis correlation coefficient is 0.72, which is statistically significant at the 99% confidence level using a boot strap resampling test. For the period 1960–1990, the average annual intensification efficiency is 14% for the tropical Atlantic, which implies the annual climatological genesis risk of a TC forming from an easterly wave in the Atlantic basin is 14%. For the period 1995–2010, the average intensification efficiency has increased to 24%, indicating that easterly waves in the present active period for TCs have a 10% greater chance of producing a tropical cyclone. This increase in the average intensification efficiency is also statistically significant at the 99% confidence level using a boot-strap resampling test.

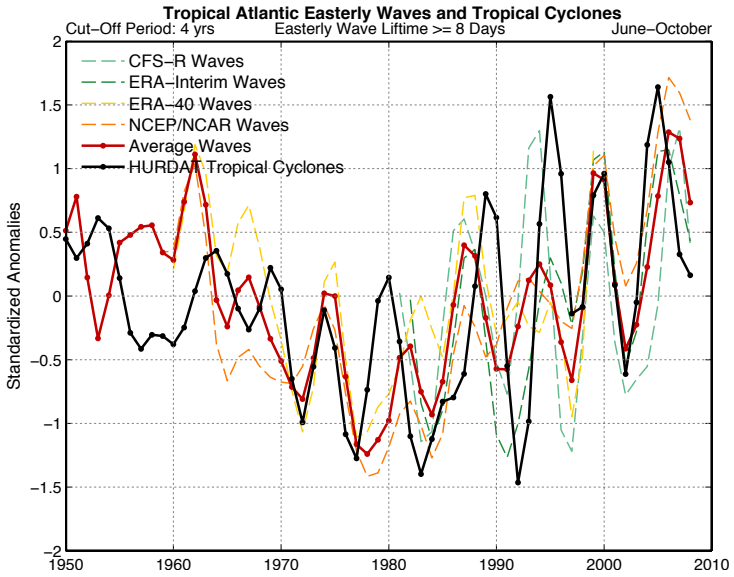

Figure 6.11: Time series of standardized anomalies of the number of easterly waves in the tropical Atlantic during the period June to October with a lifetime of at least eight days for the CFS-R (dashed light green), ERA-Interim (dashed green), ERA-40 (dashed light orange), NCEP/NCAR (dashed orange) reanalyses with the average number of waves (red) and total number of tropical cyclones (black) also included. All time series have been filtered once using a butterworth filter with three weights and a cut-off frequency of four years, and the first two/last two end-points have been excluded.

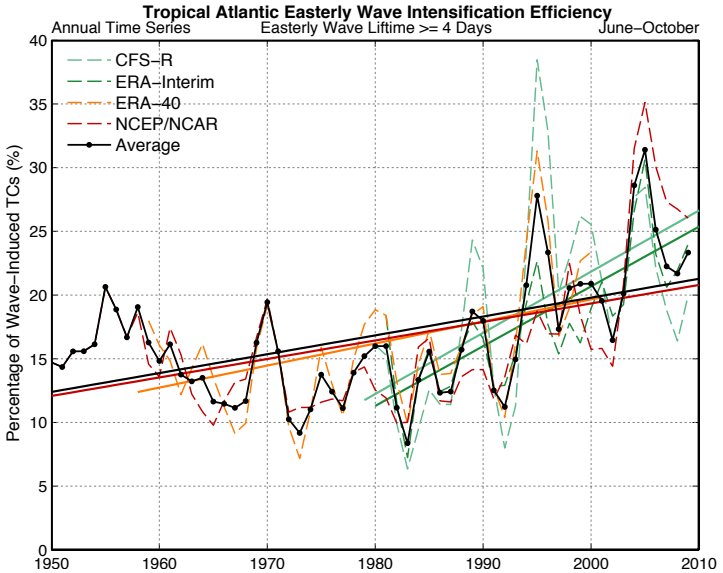

Figure 6.12: Time series of easterly wave intensification efficiency for the CFS-R (dashed light green), ERA-Interim (dashed dark green), ERA-40 (dashed orange), and NCEP/NCAR (dashed red) with the average of all four reanalyses shown in solid black with markers. All time series have been filtered once using a 1-2-1 filter. For reference, trend lines have been superimposed.

The marked increase in easterly wave intensification efficiency since the mid 1990s has important implications on the overall trend in easterly wave intensification efficiency for the period of record for each reanalysis. Table 6.4 lists the percent increase in the intensification efficiency trend for each reanalysis and as a function of easterly wave lifetime. For the periods 1948–2010 and 1958–2001, the NCEP/NCAR and ERA-40 reanalyses indicate an increasing trend of 9% and 8%, respectively, which is statistically significant at the 95% confidence level. For the periods 1979–2010 and 1980–2010, the CFS-R and ERA-Interim reanalyses have a much larger trend of 16% and 14% (also statistically significant at the 95% confidence level). Thus, it seems there are at least two dynamical mechanisms to explain the historical variability in Atlantic tropical cyclone activity. The interannual variability in easterly wave frequency modulates approximately 20–30%, depending on reanalysis, of the total variance in the number of Atlantic tropical cyclones. In addition, the recent increasing trend in the number of Atlantic tropical cyclones is tied to increased easterly wave intensification efficiency, implying that a greater proportion of easterly waves are leading to tropical cyclone formation. The underlining reason for the change in easterly wave intensification efficiency may be gleaned from recent work by Agudelo et al. (2011) who show that dynamic and thermodynamic forcing from the large-scale environment modulates up to 50% of the total interannual variability in easterly wave intensification efficiency. The large-scale environmental factors that explained the most variability include column integrated diabatic heating, mid-level specific humidity, deep-layer vertical wind shear, and sea surface temperatures.

Table 6.4: Atlantic easterly wave intensification efficiency as a function of reanalysis dataset and easterly wave lifetime for the period June to October for each reanalyses respective period of record. Easterly wave intensification efficiency is defined as the percentage of easterly waves that lead to a tropical cyclone relative to the total number of easterly waves. Values statistical significant at the 95% (99%) confidence level are in bold (bold-underline) and were determined using a boot-strap resampling test.

Intensification Efficiency Trend	Easterly Wave Lifetime ≥ 4 Days	Easterly Wave Lifetime ≥ 6 Days	Easterly Wave Lifetime ≥ 8 Days
ERA-Interim (1980-2010)	**14.1%**	**13.7%**	**14.6%**
CFS-R (1979-2010)	**14.9%**	**14.6%**	**18.3%**
ERA-40 (1958-2001)	**7.4%**	7.4%	**9.1%**
NCEP/NCAR (1948-2010)	**<u>9.0%</u>**	**8.8%**	**9.3%**

6.3.4. Easterly Waves and Climate Modes of Variability

The climatology of easterly waves can be used to understand how changes on interannual time scales may be connected to known climate modes of atmospheric and oceanic variability. First, we consider the relationship between easterly wave frequency and the Atlantic Multidecadal Oscillation index (Table 6.5). With the exception of CFS-R, the easterly wave datasets constructed from the other three reanalyses indicates a moderate, positive relationship with the AMO (correlation coefficients between 0.4 and 0.6), with the most pronounced agreement occurring with the ERA-Interim and NCEP/NCAR reanalyses. The positive correlation coefficient implies that when North Atlantic sea surface temperatures are warmer than normal, there tends to be a greater number of easterly waves that move through the tropical Atlantic basin. A similar relationship is also seen between the number of tropical Atlantic easterly waves and the Atlantic Meridional Mode (AMM). However, the similarity is not too surprising given the strong covariability that exists between the AMO and AMM (correlation of 0.79 for the period 1948–2010).

How the easterly wave frequency–AMO relationship is translated spatially may be considered using track density anomaly maps that are composite differences between positive and negative AMO phases using a standardized anomaly threshold of ± 0.5 (Figure 6.13). In this case, the anomalies have been filtered to only indicate regions that are statistically different from zero at the 90% confidence level. Qualitatively, all reanalyses indicate enhanced easterly wave track density anomalies over Africa when the AMO is in a positive phase versus its negative phase. For the ERA-Interim and ERA-40 reanalyses, both datasets show enhanced easterly wave activity in the western Sahel (west of Lake Chad). Both the CFS-R and NCEP/NCAR reanalyses show enhanced easterly wave activity mainly in the southern-tracked easterly wave zone close to the Gulf of Guinea. Across the eastern tropical Atlantic, there is some evidence to suggest a northward shift in positive easterly track density anomalies especially with the NCEP/NCAR and ERA-Interim reanalyses. However, the largest and most significant difference between positive and negative AMO phases is in the increased easterly wave frequency by 20–40%, depending on reanalysis, across the region 10^{o}N–25^{o}N x 40^{o}W–20^{o}W. Farther downstream in the tropical Atlantic the impact of AMO phasing on easterly wave track density is unclear, as the ERA-Interim and ERA-40 show higher wave activity by about 10–20%, whereas CFS-R and NCEP/NCAR both indicate reduced wave activity by 15–30%. Finally, in the East Pacific, there is some evidence to suggest that phasing of the AMO has remote impacts in terms of easterly waves that originate from the tropical Atlantic. With the exception of the CFS-R, the reanalyses show slightly higher easterly wave activity from the Atlantic moving through the East Pacific by 15% to 30%, again depending on reanalysis.

Table 6.5: Correlation coefficients between the number of easterly waves as a function of lifetime in the tropical Atlantic basin and various climate modes including the: Atlantic Multidecadal Oscillation (AMO), Atlantic Meridional Mode (AMM), North Atlantic Oscillation (NAO), Pacific Decadal Oscillation (PDO), and Nino 3.4 index for the period June to October. Values in bold are statistically significant at the 95% confidence level using a boot-strap resampling technique. The first column, Raw, denotes the annual correlation coefficient whereas the second column, Fil, denotes the filtered correlation coefficient using a butterworth filter with three weights and a cut-off period of four years.

Correlation Coefficient	AMO		AMM		NAO		PDO		Nino 3.4	
	Raw	Fil	Raw	Fil	Raw	Fil	Raw	Fil	Raw	Fil
ERA-Int.	**0.53**	**0.74**	0.34	**0.62**	-0.32	**-0.58**	**-0.39**	**-0.49**	-0.17	-0.15
CFS-R	0.15	0.11	0.03	0.00	-0.30	**-0.42**	-0.15	-0.34	-0.22	-0.28
ERA-40	0.20	0.31	0.08	0.25	-0.21	-0.21	-0.19	**-0.31**	-0.08	-0.12
NCEP	**0.54**	**0.63**	**0.34**	**0.41**	**-0.35**	**-0.34**	-0.23	**-0.37**	-0.07	-0.03

Figure 6.14 evaluates variations in the structure of 700 hPa relative vorticity and winds between positive and negative phases of the AMO. The purpose of this comparison is to provide a cursory perspective of the differences in the background climate state and should not be considered encompassing of all dynamical differences. Over Africa, the north-south dipole structure of mid-tropospheric relative vorticity exhibits a stronger gradient reversal in western Africa in positive versus negative AMO years. This pattern would favor African easterly wave amplification just prior to entry into the tropical Atlantic basin and is consistent with the theoretical findings from Leroux and Hall (2009) that describe the role the base state plays in maximizing easterly wave amplitude. Although the reanalyses differ in the orientation and longitudinal extent of the mid-tropospheric extension of the African easterly jet into the eastern Atlantic, generally the jet is weaker and displaced northward in positive AMO years versus negative AMO years among the reanalyses. In addition, in negative AMO years the dipole structure of vorticity is more elongated and has a slight positive tilt with longitude, whereas in positive AMO years, the jet structure is oriented more zonally through the eastern tropical

Atlantic. In the East Pacific, the north-south dipole of relative vorticity is similarly shifted northward and the gradient is larger than in negative AMO years. This northward shift in the wave guide may explain why positive track density anomalies are seen in both the tropical Atlantic and East Pacific.

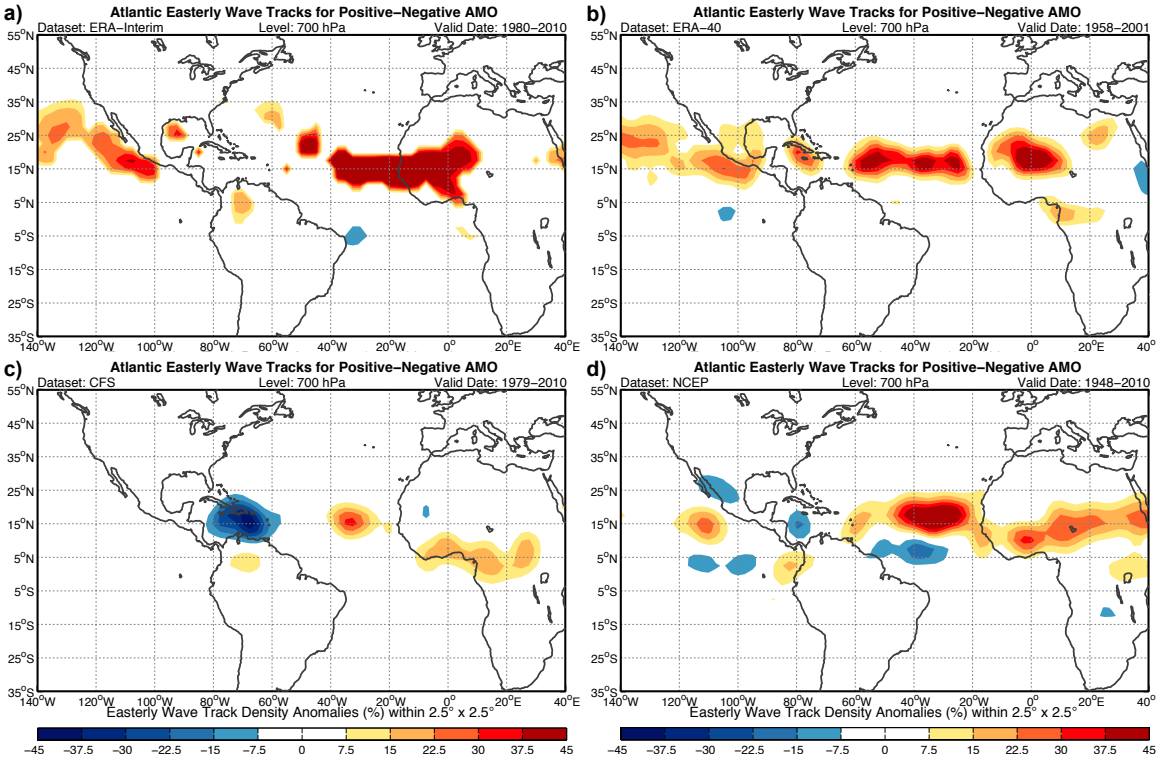

Figure 6.13: Composite Atlantic easterly wave track density anomalies between positive and negative phases of the Atlantic Multidecadal Oscillation for June to October in the period a) 1980-2010 (ERA-Interim), b) 1958-2001 (ERA-40), c) 1979-2010 (CFS-R), and d) 1948-2010 (NCEP/NCAR). Shaded regions are statistically significant at the 90% confidence level using a boot-strap resampling test. The standardized anomaly of ±0.5 in the AMO index was used to delineate between positive and negative phases.

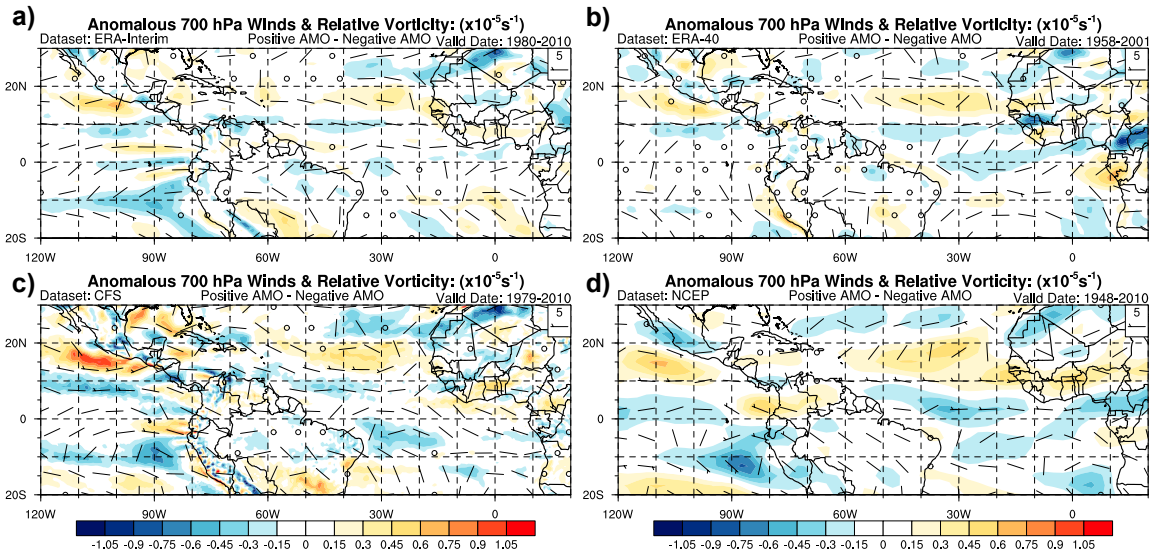

Figure 6.14: Composite anomalies of 700 hPa winds and relative vorticity after taking the difference between positive and negative phases of the Atlantic Multidecadal Oscillation for June to October in the period a) 1980–2010 (ERA-Interim), b) 1958–2001 (ERA-40), c) 1979–2010 (CFS-R), and d) 1948–2010 (NCEP/NCAR). The standardized anomaly of ± 0.5 in the AMO index was used to delineate between positive and negative phases.

The relationship between easterly wave frequency and other modes of climate variability is now considered. Table 6.5 shows the correlation coefficient between the number of easterly waves and the North Atlantic Oscillation (NAO) and Pacific Decadal Oscillation (PDO). The relationship with ENSO is not discussed, as the correlation coefficient is weak and not statistically significant at the 95% confidence level for any of the reanalysis products. With the exception of ERA-40, the other reanalyses indicate a moderate statistically, significant negative relationship at the 95% confidence level between the NAO and easterly wave frequency (correlation coefficients between -0.3 and -0.6), although the relationship is sensitive to the imposed lifetime threshold on easterly waves. The relationship also appears to be slightly stronger at periods longer than four years. Physically, the NAO may act to reduce the frequency of easterly waves by modifying the configuration of the mid-tropospheric easterly jet or through remote forcing from the North Atlantic storm track. Recent work by Chauvin et al. (2010) and Leroux et al. (2011) lends credence to the idea that the extratropics may influence

easterly wave genesis over Africa. In fact, idealized dry simulations forced with intraseasonal variability from the North Atlantic storm track are as sensitive at triggering easterly wave genesis over Africa as localized convective heating. These results provide further evidence that the mechanism for African easterly wave genesis and intensification is more complex than the baroclinic-barotropic instability mechanism originally posited by Burpee (1972).

The PDO also exhibits moderate negative covariability with easterly wave frequency across the tropical Atlantic in all reanalyses except in the CFS-R. The relationship is also spectrally greatest and most significant at periods larger than four years. Physically, it is not clear how the PDO influences easterly wave frequency in the tropical Atlantic but it may be due to a similar mechanism as the NAO through variability in the North Atlantic storm track forced by transient eddies emanating from the eastern Pacific (e.g., Trenberth and Hurrell 1994, Riviere and Orlanski 2007). However, given the weak relationship between ENSO and easterly wave frequency, it is less likely that the PDO relationship is manifested through a tropical pathway via ENSO variability.

In Section 0, the variability in easterly wave intensification efficiency for the four reanalyses was considered. The present active period for North Atlantic tropical cyclones since 1995 has seen a 10% increase in the intensification efficiency of easterly waves relative to the period from 1960 to 1990. In addition, all reanalysis products indicate a statistically significant trend in intensification efficiency over their respective period of record, which provides a dynamical explanation why the frequency of Atlantic TCs has increased during this period. Given the relationship between easterly wave frequency and various climate modes of variability including the AMO/AMM, NAO, and PDO, a logical proposition to consider is whether these climate modes are also impacting easterly wave intensification efficiency. Table 6.6 summarizes this relationship using the correlation coefficient between annually-averaged easterly wave intensification efficiency and the annual climate indices averaged for the monthly period of June to

October. The strongest relationship is found with the AMM/AMO where the relationship is moderately positive and statistically significant at the 95% confidence level for all reanalyses. The increasing trend that is observed in easterly wave intensification efficiency since 1948 or since 1980 is not statistically different from the increasing trend that has been observed with the AMO (not shown). In short, regressions of the AMO index onto easterly wave intensification efficiency entirely capture the observed increasing trend in easterly wave intensification efficiency for each reanalysis. There is some evidence of a statistically, significant negative relationship between easterly wave intensification and both the NAO and PDO as indicated with the NCEP/NCAR and CFS reanalyses, respectively. In addition, unlike the relationship between ENSO and easterly wave frequency, a moderate negative relationship between the Nino 3.4 index and easterly wave intensification efficiency is seen and is spectrally significant at periods longer than four years for the CFS-R and NCEP/NCAR reanalyses.

Table 6.6: As in Table 6.5, except for the average correlation coefficient between easterly wave intensification efficiency and various climate modes.

Correlation Coefficient	AMO		AMM		NAO		PDO		Nino 3.4	
	Raw	Fil	Raw	Fil	Raw	Fil	Raw	Fil	Raw	Fil
ERA-Int.	**0.56**	**0.69**	**0.66**	**0.75**	-0.29	-0.13	-0.35	-0.38	-0.02	-0.35
CFS-R	**0.56**	**0.62**	**0.58**	**0.69**	-0.07	0.18	**-0.43**	**-0.42**	-0.21	**-0.43**
ERA-40	**0.45**	**0.41**	**0.51**	**0.47**	-0.08	0.04	-0.06	0.13	-0.09	-0.28
NCEP	**0.54**	**0.64**	**0.58**	**0.59**	**-0.31**	**-0.39**	-0.18	-0.08	0.04	**-0.26**

6.4. Conclusions

To understand the relationship between easterly waves, tropical cyclones and the large-scale environment on interannual time scales, a robust climatology of easterly waves for the tropical Atlantic and East Pacific has been developed. The new easterly

wave-tracking scheme reproduces westward-propagating disturbances with structures and phase speeds characteristic of easterly waves over Africa and in the East Pacific. More importantly, the new scheme eliminates some of the deficiencies found in previous tracking algorithms (i.e. Hodges et al. 1995, Agudelo et al. 2011) in terms of locating easterly waves that move across Central America and into the East Pacific as well as ensuring the westward propagating disturbances are mostly synoptic-scale easterly waves and not mesoscale convective systems.

A comparison of the easterly wave track climatologies developed for the ERA-Interim, ERA-40, CFS-R, and NCEP/NCAR reanalyses reveals a number of important similarities and differences. All easterly wave climatologies reproduce the seasonal cycle in terms of the preferred frequency of easterly waves for both the East Pacific and tropical Atlantic. Easterly waves are generated in both the East Pacific and tropical Atlantic during all months of the year. However, these structures would go undetected if only outgoing longwave radiation was used for tracking. The NCEP/NCAR reanalysis produces a significantly larger number of easterly waves overall for both basins while the other reanalyses show better agreement in terms of the seasonal climatology. Composite maps of easterly wave track density for the tropical Atlantic and the East Pacific also provide unique information. All reanalyses indicate that there are two preferred easterly wave genesis regions both north and south of the equator during the period May to October. Generally, a significantly greater number of easterly waves are generated in the summer hemisphere with easterly waves generated north of the equator having the largest overall lifetime and reaching the most westward longitudinal position. Both the ERA-40 and NCEP/NCAR reanalyses generally favor easterly wave genesis farther east across northern Africa than either the CFS-R or ERA-Interim datasets. These older reanalyses also indicate a greater proportion of easterly waves moving into the East Pacific that are generated over Africa or in the eastern tropical Atlantic than more recent global reanalyses. In terms of easterly waves that move through the East Pacific, there is a

similar split in terms of preferred trajectories for easterly waves both north and south of the equator during the boreal summer. However, relative to the trajectory of easterly waves originating over Africa or the tropical Atlantic, there is an equatorward shift in the latitudinal location of easterly wave formation in the East Pacific. Many of these differences in terms of preferred zones for easterly wave genesis, maximum amplitude, or seasonal frequency are predicated on the variability in the climatological structure and strength of the mid-tropospheric easterly jet and the location of the maximum meridional gradient reversal in relative vorticity.

Using the derived easterly wave climatologies for each reanalysis, the covariability between easterly wave frequency and intensification efficiency is also examined relative to the annual number of tropical cyclones in the North Atlantic. A moderately positive statistically significant relationship is seen with all reanalyses except with the CFS-R. This relationship is sensitive to the imposed easterly wave lifetime threshold likely due to biases in the preferred region of easterly wave genesis over Africa. Generally this relationship is strongest at periods beyond four years. In fact, cospectrum and coherence squared analyses of the two time series indicates power at four to five year periods and also at 20-year periods, although the relationship at these longer time scales was not deemed statistically significant perhaps due to the limited length of the easterly wave climatology.

The lack of relationship between easterly wave frequency and TCs in the CFS-R along with differences in the easterly wave track density suggests that the easterly wave tracking thresholds for the CFS-R may need to be revised. Even though the same procedure is used to decimate its high spatial resolution to a coarse 2.5^{o} x 2.5^{o} grid, the 6-hr variance in curvature vorticity on this coarse grid is still significantly larger than any other reanalysis (c.f. Figure 6.3c with Figure 6.3a,b,d). A topic of future work is to examine easterly wave sensitivity in the CFS-R using larger curvature vorticity tracking thresholds.

The relationship between easterly wave frequency and various climate modes including the Atlantic Multidecadal Oscillation, Atlantic Meridional Mode, North Atlantic Oscillation, Pacific Decadal Oscillation, and El Nino Southern Oscillation is also considered. Both the AMO and AMM exhibit the strongest positive covariability with Atlantic easterly wave frequency. Spatial track density anomalies between positive and negative AMO years indicates enhanced easterly wave genesis across the western Sahel, larger track density anomalies across the tropical Atlantic's MDR, and some suggestion of a northward shift in preferred zones of easterly wave tracks in the tropical Atlantic. The magnitude of the relationship between easterly wave frequency and the NAO and PDO is weaker than with the AMO, and the relationship is inversely proportional, which implies that during positive NAO/PDO years, the number of easterly waves across the tropical Atlantic is on average lower than during negative NAO/PDO years. The hypothesis that has been offered to explain how the NAO and PDO modulate easterly wave frequency in the tropical Atlantic is through induced changes to the North Atlantic storm track, which has recently been shown to be as sensitive to localized convective heating over Africa at triggering easterly wave genesis (Leroux et al. 2011). The NAO may also modulate the strength and structure of the mid-tropospheric extension of the AEJ into the tropical Atlantic, although both these possible mechanisms remain a topic of future work.

The intensification efficiency of easterly waves is shown to have substantial interannual variability. The coefficient of variation, which is the ratio of the standard deviation to the mean, averages 50% for the four reanalyses over their respective period of record. The intensification efficiency of easterly waves has increased during the recent active period of tropical cyclones when compared to earlier periods. Previously, the average intensification efficiency was 14% during the period 1960–1990, but since 1995 the intensification efficiency has increased by 10%. Now, there is a 24% probability that a given easterly wave in the tropical Atlantic will induce a tropical cyclone at some point

during its period of movement across the basin. When variability in easterly wave intensification efficiency is considered relative to various large-scale climate modes, an even stronger relationship is seen with the AMO and AMM than was seen with easterly wave frequency. In fact, the relationship is so robust that regressions of the AMO onto easterly wave frequency entirely reproduce the low-frequency trends that have been observed in both the ERA-Interim and CFS-R since 1980 and the NCEP/NCAR and ERA-40 since 1948 and 1958, respectively.

The relationship between easterly waves and tropical cyclones on interannual time scales has revealed two dynamical mechanisms that may explain how the frequency of Atlantic tropical cyclones has varied with time. Interannual variations in the number of easterly waves modulate approximately 20–30% of the total variance in the annual number of tropical cyclones. In addition, the increasing trend in the number of Atlantic tropical cyclones is concomitant with the seasonal number of easterly waves. Although the number of easterly waves has varied interannually and shows the same increasing trend as tropical cyclones, we also find large interannual variability with a similar increasing trend in the intensification efficiency of easterly waves. This finding suggests that in spite of the variability in the number of easterly waves, the ratio of induced tropical cyclones from easterly waves relative to the total number of waves in a given season is not fixed but evolves with time. In this respect, dynamic as well as thermodynamic factors from the large-scale environment such as column integrated diabatic heating, vertical wind shear, sea surface temperatures, and mid-level specific humidity may largely account for the interannual variability in easterly wave intensification efficiency as shown in Agudelo et al. (2011).

CHAPTER 7
INTERANNUAL VARIABILITY OF U.S. LANDFALLING TROPICAL CYCLONES AND ITS IMPACT ON TC TORNADOES

The work presented in this chapter is published in Geophysical Research Letters (Belanger et al. 2009).

7.1. Background and Motivation

The recent increase in North Atlantic tropical cyclone (TC) frequency has caused a concomitant rise in the number of landfalling TCs especially along the Gulf of Mexico coast that has reached or possibly exceeded previous periods of elevated activity (Goldenberg et al. 2001). Although the impact from this variability in U.S. landfall frequency and from tropical cyclones in general may come in many forms, the focus of this chapter is on one, often-overlooked impact from TCs that can be particularly destructive for inland regions: TC-induced tornadoes.

As discussed in Section 1.4, since the identification of a TC tornado is determined by a variety of sources, the climatology of TC tornado reports has a large undercounting bias especially prior to National Weather Service Modernization in the mid-1990s. To avoid artificially correcting the inhomogeneous TC tornado record, a statistical model is developed for the tornado frequency produced by a landfalling TC using a set of a priori predictors for TC tornado formation including: TC size, intensity, mid-level specific humidity gradient, and TC recurvature. The statistical model is trained on the period 1998–2007 as the national network of NEXRAD (WSR-88D) weather radars was completed in 1997 and reflects a period of near homogeneous observational detection for tornadoes (Simmons and Sutter 2005). The utility of this model for TC tornadoes lies in

its capability as a real-time, forecast tool as well as in reconstructing the climatology of TC tornadoes in the U.S.

This chapter is divided into the following sections. The data and methodology to build the statistical TC tornado models and to reconstruct the historical TC tornado record is presented in Section 7.2. In Section 7.3 the statistical tornado models are verified using the training period 1998–2007 and an independent evaluation of the models is conducted using the 2008 Atlantic hurricane season, which featured five Gulf landfalling TCs. These results are followed by a discussion of the reconstructed interannual variability in the number of TC tornadoes and its relationship with historical changes in the number of U.S. landfalling TCs. Finally, concluding remarks are provided in Section 7.4.

7.2. Data and Methodology

The Hurricane Database for the North Atlantic was used to highlight U.S. landfalling TCs from 1920 to 2008. Maximum TC intensity at landfall was taken from HURDAT using the maximum 1-minute sustained wind speed, and based on the tracks for U.S. landfalling TCs, these systems were stratified into two groups: Gulf TCs and Atlantic TCs. For TCs with multiple landfalls such as Hurricane Katrina of 2005, these TCs were considered Gulf landfalls if the system made landfall at some point along the Gulf coast. Here, TC tornado statistics are examined for Gulf TC landfalls only, which account for 85% of all reported TC tornadoes (e.g., McCaul 1991). Using HURDAT and the Gulf TC landfall dataset, TC recurvature is calculated as the difference in 12-hr averaged storm heading at landfall from the 12-hr averaged storm heading at the last NHC tropical advisory or the last tropical advisory while over the U.S. mainland, whichever occurs first.

The National Climatic Data Center's (NCDC) Storm Events Database is used to identify all tornadoes that occurred in the U.S. prior to and after a TC's landfall during

the period 1950–2008. Since tornado entries in the Storm Events Database are divided into county segments, tornado listings were compared to one another to ensure that no tornado was counted more than once. A TC tornado is defined to be any tornado that formed within 650 km of the center of circulation before the last tropical advisory was issued for a TC. Note: McCaul's (1991) database uses 800 km as the distance threshold. In addition, surface weather maps at the time of landfall were examined to eliminate tornadoes that developed after the TC had undergone extratropical transition or those that may have formed due in part to short-wave trough interactions or other mechanisms not related to the landfalling TC. With this information, a TC tornado database was created from 1950 to 2007 and is available in Belanger et al. (2009). Comparing McCaul's TC tornado database with this study reveals slight differences in the total number of tornadoes for some landfalling tropical cyclones, which are mainly attributed to the inclusion of tornadoes that are likely associated with TCs that have undergone extratropical transition (not due to differences in the tornado distance threshold).

A key predictor for the number of TC tornadoes is the horizontal size of the TC at landfall. Measures of TC size include: radius of the eye, radius of maximum winds, radius of hurricane force winds, radius of gale force winds, and radius of outer closed isobar (Kimball and Mulekar 2004). In this analysis, the radius of the outer closed isobar (ROCI) is used as a proxy for TC size, since it is more closely related to the domain over which TC tornadoes may occur and since the ROCI analysis can be extended into the pre-satellite era. Following Merrill (1984), the ROCI best estimate is obtained by averaging the distances (in km) to the north, south, east, and west from the cyclone center to the closed isobar having the highest pressure. In this study the landfall ROCI database was constructed using archived daily weather maps from the NOAA Central Library Data Imaging Project near the time of landfall for each TC from 1920 to 2001. For TC landfalls occurring after 2001, archived surface weather maps from NCDC's Service Records Retention System were used. The surface weather maps contoured surface

pressure at intervals of 0.1 in of Hg from 1920 to 1939 and at 4 hPa intervals from 1940 to present. For cases when the outer closed isobar was elongated or distorted, the second highest value closed isobar was used as the outer closed isobar. For cases when the isobars around a landfalling tropical cyclone were influenced by synoptic-scale features such as a mid-latitude cyclone or a frontal boundary, these TCs were excluded from the ROCI analysis. Note: A total of 203 TCs have made landfall from the Gulf of Mexico since 1920, and 185 were included in the ROCI analysis.

The final predictor variable is the specific humidity gradient at 500 hPa (within 5^{o} latitude-longitude from the TC center) obtained from the NCEP/NCAR Reanalysis I dataset. It is hypothesized that the 500-hPa specific humidity gradient represents a measure for mid-level dry air entrainment. Physically, dry air intrusions in a TC's outer-rainbands induce convective downdrafts that then establish localized baroclinic boundaries in an otherwise moist, lower troposphere. These cold pools, which may be a source of horizontal vorticity (McCaul and Weisman 1996), act in conjunction with enhanced low-level wind shear within the right-front quadrant of a landfalling TC to increase the probability of a tornado outbreak (Gentry 1983, McCaul 1991, Curtis 2004).

7.2.1. Development of a Statistical TC Tornado Prediction Model

A statistical model for the TC tornado frequency per Gulf TC landfall is developed from Equation 7.1

$$y = a_0 + a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4 + \varepsilon \qquad 7.1$$

where y is TC tornado frequency and x_1, x_2, x_3, and x_4 are ROCI, maximum wind speed at 10 m, 500 hPa specific humidity gradient, and TC recurvature, respectively. Integrated and scaled probability density functions (PDFs) for each predictor and TC tornado frequency were used to determine the power of each predictor variable.

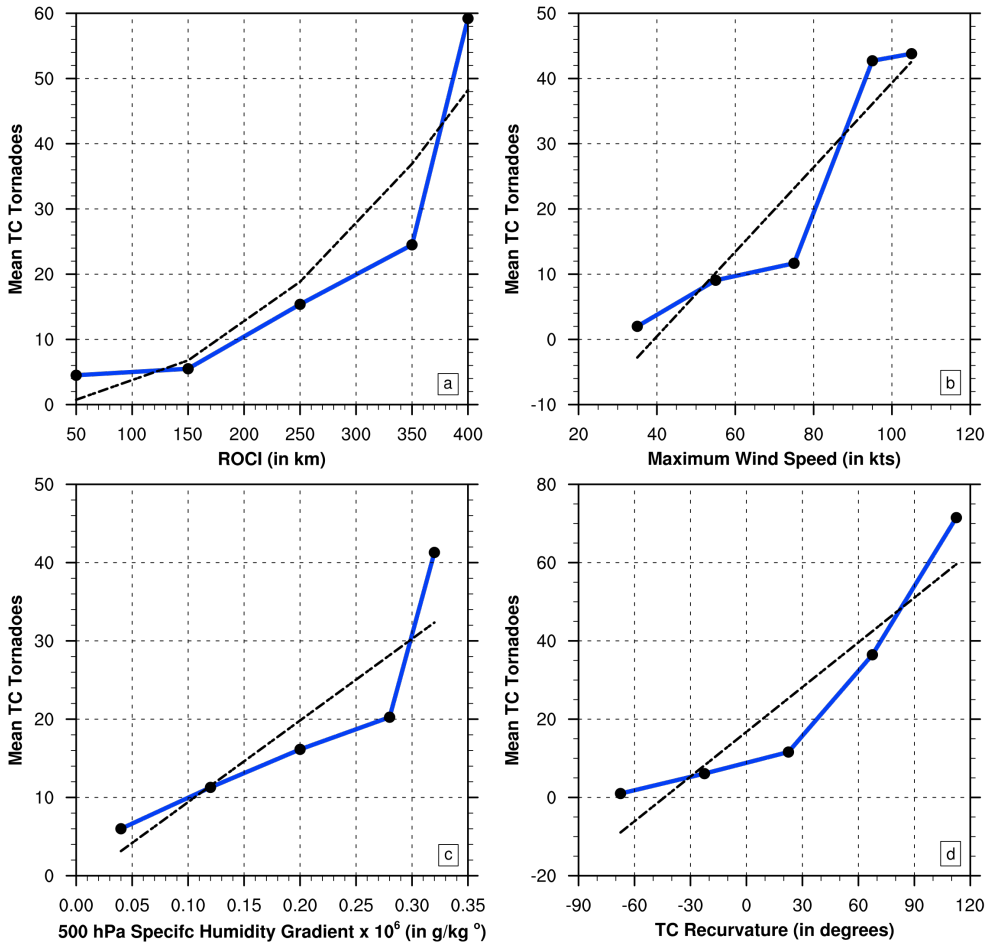

Figure 7.1: a) Integrated and scaled probability density functions for the mean number of TC tornadoes as a function of radius of outer closed isobar observations for all Gulf landfalling TCs in the period 1990–2008 (blue) along with a second order polynomial (in dashed black). b) Similar to a), except for the mean number of TC tornadoes as a function of maximum 1-minute sustained wind speed at 10m (in kt) at landfall along with a first order polynomial (in dashed black). c) Similar to a), except for the mean number of TC tornadoes as a function of 500 hPa specific humidity gradient at 5° latitude-longitude from the TC center along with a first order polynomial (in dashed black). d) Similar to a) except for the mean number of TC tornadoes as a function of TC recurvature along with a first order polynomial (in dashed black).

Figure 7.1a reveals that the relationship between the average number of TC tornadoes and ROCI is parabolic and is best fitted with a second order polynomial, indicating that the average number of TC tornadoes is proportional to the square of ROCI. Since the horizontal area impacted by an axisymmetric tropical cyclone is proportional to the squared radius of the TC, this result is physically consistent with the expectation that larger Gulf TCs are more likely to produce a larger number of tornadoes and is in agreement previous work by McCaul (1991). Figure 7.1b-d show that the relationship between the number of TC tornadoes and the other predictors (maximum wind speed, 500 hPa mid-level specific humidity gradient at 5^{o} latitude-longitude from the TC center at landfall, and TC recurvature) is most likely linear. Underlining each predictor's relationship is that the standard deviation of the mean number of TC tornadoes increases nonlinearly as the predictor increases, which suggests that a variety of processes and variables, likely acting in conjunction, are important in determining whether a large tornado outbreak will occur as a TC makes landfall along the Gulf coast.

In Equation 7.1, the ε term describes the model residuals, and the a_s terms are regression coefficients that were determined from the following operation. If the model residuals are ignored, then this equation may be written into the one below (Equation 7.2).

$$\mathbf{y}_n = \mathbf{X}\mathbf{A}_s \tag{7.2}$$

Equation 7.2 may be solved for the regression coefficients vector, which is provided in Equation 7.3.

$$\mathbf{A_s} = \left(\mathbf{X'X}\right)^{-1}\mathbf{X'y_n} \tag{7.3}$$

Two separate multiple regression models are developed: Recon, which employs all four predictor variables, and Extended Recon, which uses TC intensity, size, and recurvature as predictor variables. Both models are trained on the 1998–2007 period of observed TC tornado data and Recon (Extended Recon) is used to reconstruct the TC

tornado climatology back to 1948 (1920). Statistical significance tests are used to assess whether changes in the reconstructed TC tornado data set are due to random chance. The 95% confidence level is determined by bootstrapping the Recon and Extended Recon model residuals for the period 1998–2007.

7.3. Results

7.3.1. Assessment of Statistical Models

To evaluate the Recon and Extended Recon models, Figure 7.2 shows a scatter plot of the observed and modeled tornado frequency for each Gulf landfalling TC during the period 1998–2007. Overall, the Recon and Extended Recon models explain 70% and 62%, respectively, of the total variance in the number of observed tornadoes per Gulf TC landfall from 1998 to 2008 (Table 7.1). Both models have the largest errors for TCs with the lowest and highest number of observed TC tornadoes, which is characteristic of least square statistic methods (Wunsch 2006). In the Recon model, the ROCI predictor explains the most variance in the observed number of tornadoes per Gulf TC landfall at 43%, followed by the 500 hPa specific humidity gradient (17%), maximum wind speed (8%), and TC recurvature (2%). For the Extended Recon model, ROCI explains 43% of the variance in the observed frequency of TC tornadoes with TC recurvature and TC intensity accounting for 11% and 8%, respectively. For the annual number of TC tornadoes during the 1998–2008 period, both models performed very well in capturing the seasonal total of Gulf TC tornadoes as the variance explained by the Recon and Extended Recon models is 90% and 85%, respectively.

An independent evaluation of both models was undertaken by applying them to the 2008 Atlantic hurricane season, which featured a record six consecutive TC landfalls on the U.S. mainland, five of which were Gulf TCs (Table 7.2). For four out of the five Gulf landfalling TCs from 2008, the Recon model captured the magnitude of the observed TC tornado outbreak that resulted post-landfall. Given the large mid-level

specific humidity gradient at landfall for Tropical Storm Fay, Recon predicted an outbreak of 45 TC tornadoes would occur in comparison to a reported 44 TC tornadoes that ensued as Fay slowly moved inland. The accuracy in the TC tornado forecast for Hurricane Dolly is attributed to the TC recurvature predictor as Dolly did not recurve to the northwest post-landfall but maintained a west-northwest storm heading. Overall, the Recon and Extended Recon models predicted a seasonal total of 123 and 98 TC tornadoes, respectively, in comparison to an observed value of 109 TC tornadoes.

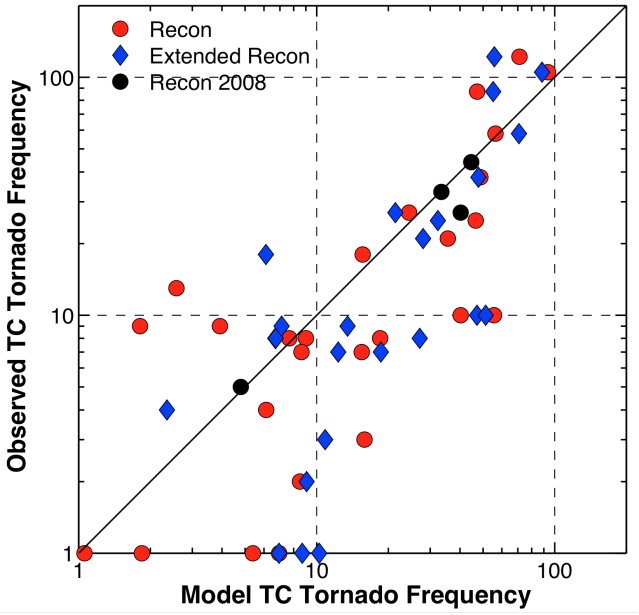

Figure 7.2: Log-log scatter plot of observed versus modeled tornado frequency per Gulf TC from the Recon (red circles) and Extended Recon (blue diamonds) models for the period 1990–2007. Black circles denote the observed versus modeled TC tornado frequency from Recon for 2008.

Table 7.1: Performance of the Recon and Extended Recon models in terms of the correlation coefficient and root mean square (RMS) error for the 1998–2008 period.

1998–2008	Correlation (TC)	RMS Error (TC)	Correlation (Annual)	RMS Error (Annual)
Recon	0.84	16	0.95	26
Extended Recon	0.79	18	0.92	30

Table 7.2: TC tornado statistics for the five Gulf landfalling TCs from the 2008 hurricane season.

Tornadoes 2008 Season	Observed	Recon	Extended Recon
HR Dolly	5	5	0
TS Edouard	0	0	0
TS Fay	44	45	18
HR Gustav	27	40	44
HR Ike	33	33	36
Total Tornadoes	109	123	98

7.3.2. Annual Variability of TC Tornadoes

Due to a large undercounting bias throughout the historical record, the observed TC tornado dataset does not allow for an assessment of the interannual variability of tornado frequency due to Gulf landfalling TCs. Instead, examining the filtered seasonal reconstructions from Recon and Extended Recon reveals that the early 1930s and 1940s were likely associated with enhanced TC tornadic activity and the 1970s and early 1980s with reduced activity (Figure 7.3). The time period since 1995 has been marked by a high seasonal number of TC tornadoes, with 2004 and 2005 having reached unprecedented levels for annual TC tornado frequency according to the 95% confidence interval from the Recon and Extended Recon model reconstructions.

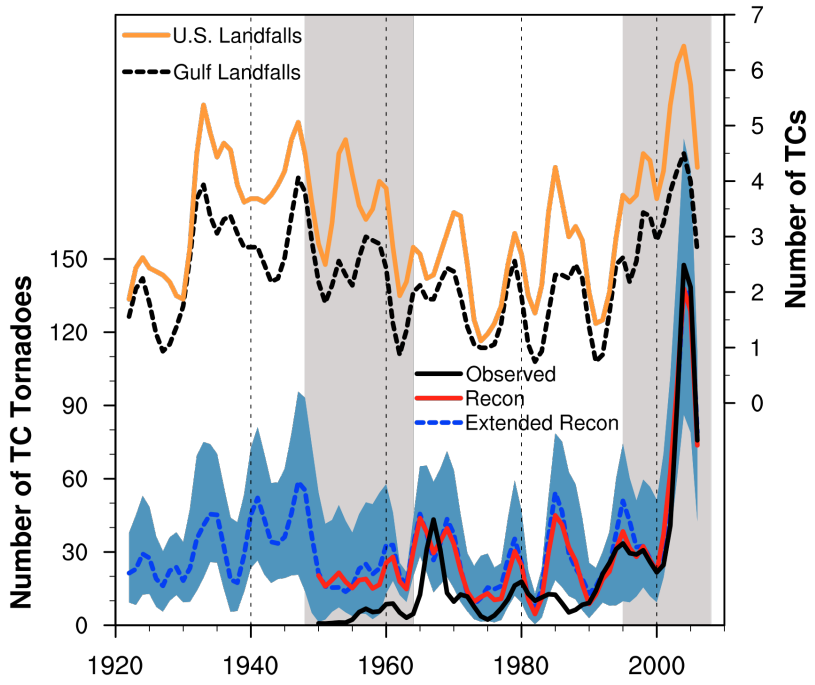

Figure 7.3: (top) Time series of the number of U.S. landfalling TCs (Gulf landfalling TCs) from 1920 to 2008 in orange (dashed black). (bottom) Time series of the annual number of TC tornadoes from Recon and Extended Recon in red and dashed blue for the periods 1948–2008 and 1920–2008, respectively, along with the observed TC tornadoes in black from 1950 to 2008. All time series have been filtered twice with a 1-2-1 filter. The 95% confidence interval is shaded in light blue, which for 1948–2008 (1920–1947) is determined by bootstrapping the Recon (Extended Recon) model residuals for the period 1998–2007. Gray boxes represent the two active periods for Gulf TCs: early active (1948–1964) and current active (1995–2008).

Both the decadal scale fluctuations in seasonal Gulf TC landfalls and TC tornadoes reflect the same multidecadal pattern of variability associated with the Atlantic Multidecadal Oscillation with warm phases of the AMO in 1926–1964 and 1995–2008 corresponding to active periods for Gulf landfalling TCs. However, relative to the previous periods of peak Gulf TC activity, the number of Gulf TC landfalls during the period 1995–2008 is not unusual. Hence, the recent elevation of TC tornadic activity cannot be attributed to an increase in the number of landfalling TCs from the Gulf of Mexico.

Since the variability in Gulf TC landfall frequency does not explain the high annual number of TC tornadoes in recent years, this result suggests that the number of tornadoes per TC landfall has changed between the two periods for Gulf landfalling TCs. To test this hypothesis, Figure 7.4a provides the PDFs for the number of TC tornadoes per Gulf TC landfall from the Recon model for the two active periods, 1948–1964 (early active) and 1995–2008 (current active). The PDF for the current active period has broadened from the early active period and has seen a statistically significant increase (95% confidence level) in the frequency of large TC tornado outbreaks. In addition, the median number of tornadoes per Gulf TC landfall has significantly increased (95% confidence level) from 6 to 15 TC tornadoes.

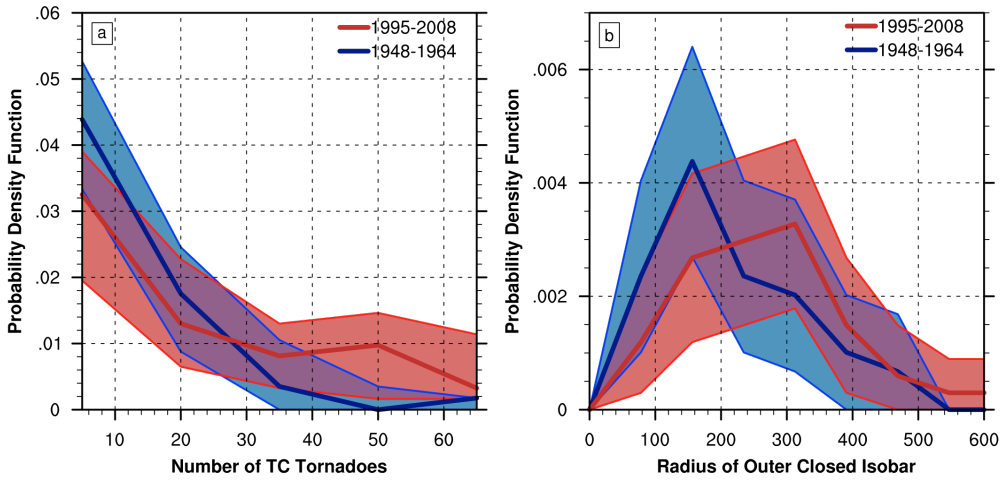

Figure 7.4: (a) Probability density functions of TC tornado frequency per Gulf TC landfall for the early active period (1948–1964) in thick blue and current active period (1995–2008) in thick red using the Recon model. The shaded regions of light blue (light red) indicate the 95% confidence interval of the early active (present active) PDF. Confidence interval determined by bootstrapping the Recon model residuals for the period 1998–2007. (b) Similar to Figure 7.4a, except for the ROCI.

Since these changes in the PDFs for TC tornado frequency were determined using the Recon model, the analysis focuses on determining which predictors are responsible for these differences. The only variable that shows a statistically significant increase (95% confidence level) in the median and extreme values and makes an important

contribution to the TC tornado predictions in the Recon model is ROCI (c.f. Figure 7.4b and Figure 7.5a,b,c). The median ROCI has increased from 195 km to 263 km between the two active periods. In addition, the frequency of large TCs (ROCI > 540 km) has significantly increased (95% confidence level) from the early active period. This change in the ROCI distribution is restricted to the years after 1995 as the PDF for Gulf landfalling TCs for the 1965–1994 period (not shown) exhibits a statistically similar structure to the 1948–1964 period in terms of median ROCI (205 km) and a lack of extremely large TC landfalls.

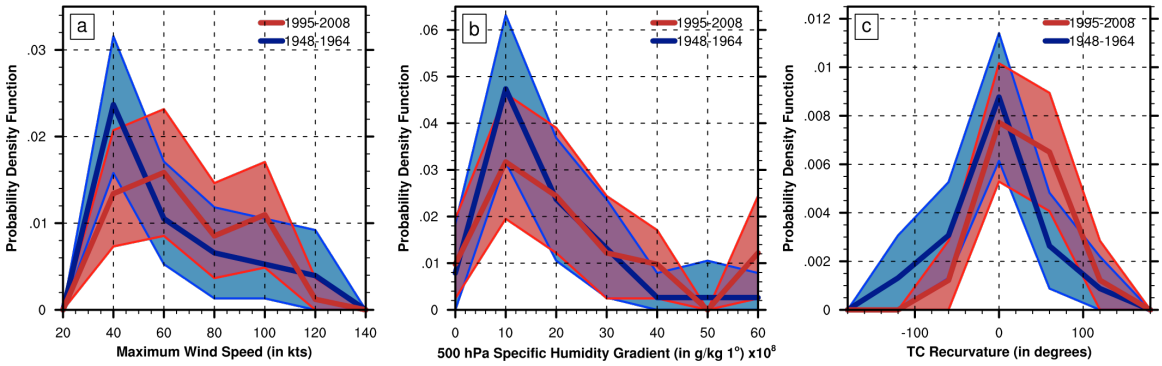

Figure 7.5: (a) Probability density functions of TC intensity for Gulf landfalling TCs from the early active period (1948–1964) in thick blue and current active period (1995–2008) in thick red. The shaded regions of light blue (light red) indicate the 95% confidence interval of the early active (present active) PDF, which was determined by bootstrapping from the TC intensity data for each period, respectively. (b) Similar to Figure 7.5a, except for the 500 hPa specific humidity gradient at 5° latitude-longitude from the TC center. (c) Similar to Figure 7.5a,b, except for TC recurvature.

To address the possibility that the change in Gulf TC size between the 1948–1964 and 1995–2008 periods is due to the subjective analysis of surface weather maps to determine a tropical cyclone's ROCI, another TC size metric is used which is the distance of tornado (DOT) from TC center. Since most tornadoes occur in the outer-rainbands of a landfalling tropical cyclone, DOT observations serve as a proxy measure for the location of outer-rainbands and ultimately the size of a landfalling TC. Figure 7.6 shows the probability density function of DOT observations from Gulf landfalling TCs from the periods 1948–1994 and 1995–2008. Also, given a lack of tornado observations

associated with Gulf landfalling TCs especially in the 1950s, the pre-1995 DOT PDF is compared to the PDF from the current active period since 1995. Consistent with the observed changes in the probability distributions of ROCI, the DOT distributions exhibit a statistically significant (99% level) increase in the median distance of tornado formation (289 km versus 336 km) and an increased frequency of tornadoes occurring at large distances from the TC center. Quantitatively, the magnitude of the increase in median TC tornado distance is smaller (by 21 km) than the median ROCI increase, but nevertheless, the qualitative results between the two measures of TC size are considered robust. These changes in DOT observations are physically consistent if the size of Gulf TCs at landfall has increased between the two active periods.

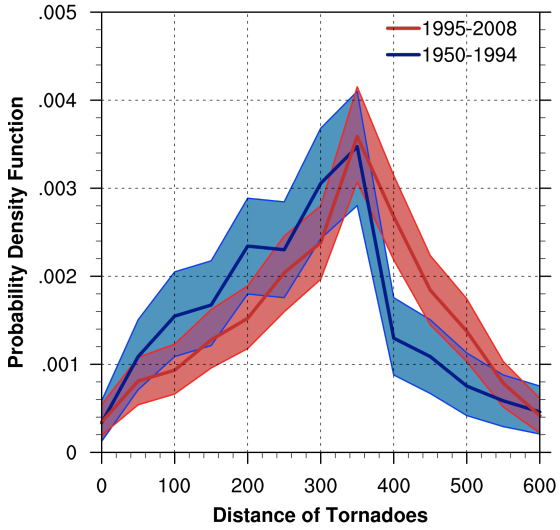

Figure 7.6: Probability density function for the distance of tornado from TC center observations for the periods 1950–1994 in thick blue and 1995–2008 in thick red. The shaded regions of light blue (light red) denote the 95% confidence interval of the 1950–1994 (1995–2008) PDF determined by bootstrapping from the DOT data for each period.

7.4. Conclusions

A new data set has been assembled for tornadoes associated with Gulf landfalling TCs that is used to develop a statistical model for TC tornadoes as a function of TC intensity, horizontal size, mid-tropospheric specific humidity gradient at landfall, and TC

recurvature. Two forms of the statistical model are developed including Recon, which includes all four predictors, and Extended Recon, which does not include specific humidity data. For the period 1998–2008, the Recon (Extended Recon) model explains 70% (62%) of the observed TC tornado frequency per Gulf TC landfall and 90% (85%) of the seasonal number of TC tornadoes. Applying the models to the 2008 Hurricane Season, the Recon and Extended Recon models predicted an annual total of 123 and 98 TC tornadoes, respectively, in comparison to 109 reported TC tornadoes.

In addition to its utility as a prediction tool for individual TCs, the Recon and Extended Recon models are used to develop a synthetic climatology of historical TC tornadoes back to 1948 and 1920, respectively. The reconstructions from the models along with the 95% confidence interval show that the high number of TC tornadoes during 2004 and 2005 is unprecedented in the historical record back to 1920. Also, the synthetic climatology of TC tornadoes clearly reflects the decadal scale variations in Gulf landfalling TC activity associated with the AMO. A comparison of the current warm phase of the AMO (1995–2008) with the previous warm phase period from 1948–1964 shows that the median number of TC tornadoes per Gulf TC landfall according to the Recon model has significantly increased (95% confidence level) from 6 to 15 TC tornadoes along with an increased frequency of large TC tornado outbreaks. PDFs of Gulf TC size show a 35% increase in median Gulf TC size along with an increased frequency of large Gulf landfalling TCs. These changes in TC size explain why the Recon model projects a significant increase in the median and extreme frequency of tornado events between the two active periods of Gulf landfalling TCs.

The increase in TC size during the current active phase of North Atlantic TC activity raises the possibility that climate variability is modulating the change in Gulf TC size. Recent research by Hill and Lackmann (2009) has shown that environmental relative humidity places an important constraint on the horizontal extent of a developing TC. We hypothesize that increasing environmental moisture in the lower troposphere

over a tropical cyclone's lifetime is responsible for the observed increase in median and/or extreme frequency of large Gulf TCs. Additional research on TC size, including environmental influences and internal dynamics, is clearly warranted for understanding and assessing their landfall impacts in the United States.

CHAPTER 8
CONCLUDING REMARKS

The continual evolution of global numerical weather prediction systems necessitates constant evaluation of the practical improvement in atmospheric predictability, especially for mesoscale phenomena whose structural characteristics become better resolved at finer spatial scales. For tropical cyclones, this process has led to the revision in forecasting practices over time for TC genesis, track movement, and intensity change. For instance, the National Hurricane Center did not begin operational TC genesis forecasts until 2008, after the research community (e.g., Pasch et al. 2006, Fontan and Cabanes 2008, etc.) had determined that the newest generations of global NWP systems had skill at predicting TC genesis for the next 48 hr. Tropical cyclone track forecasts share a similar pattern of evolution. During the late 1990s, TC track research (i.e., Aberson 1998) began to indicate that forecast skill exceeding climatology and persistence could be provided for forecasts beyond 72 hr. These findings, along with improvements in global NWP systems, eventually led the NHC to extend their forecast time horizon from three to five days in 2001.

Based on the TC predictability studies using the VarEPS in the North Indian and North Atlantic Oceans in Chapter 3 and Chapter 5, these findings provide the basis for future improvements to operational tropical cyclone forecasts. First, in Chapter 3 the VarEPS is shown to provide well-calibrated probabilistic forecasts for TC genesis through a lead-time of one week, with low false alarm rates and moderate to high probabilities of detection. Coupled to these TC genesis forecasts is the ability of the VarEPS to provide pregenesis forecasts with similar or enhanced forecast skill (at lead-times beyond 48 hr) compared to the VarEPS's postgenesis track forecasts (Chapter 3).

If extended TC genesis forecasts along with pregenesis track outlooks are transferred into an operational forecast system, then the practical impact could be quite significant, especially for regions where TCs tend to form close to the coast, have short lifetimes, and hence, until now, short forecast warning windows (e.g., North Indian Ocean, Australia, North Atlantic).

Other advances that have been discussed in this dissertation include using updated statistical calibration techniques to improve global NWP track and intensity forecasts. Previous approaches (e.g., Goerss 2000, Elsberry and Carr 2000) have attempted to build multi-model consensus forecasts using a weighted linear combination of the most skillful global or regional forecast models. This approach has been examined for other prognostic variables such at low-level temperatures and 500 hPa geopotential heights through TIGGE, and initial results show that the multi-model approach often does not lead to a better forecast than calibration of the single best global model (e.g. Hagedorn et al. 2012). Furthermore, global NWP systems are evolving constantly, which requires the weighted coefficients to be recalculated after each hurricane season.

Given the sample size limitations with the consensus technique and the fact that past forecast skill is not necessarily indicative of future performance, a new approach to track and intensity calibration is utilized in Chapter 5. The technique assumes that along-track, cross-track, and wind intensity biases in the ECMWF hindcasts also characterize the real-time VarEPS. When these historical errors are used in a Monte Carlo resampling technique that removes initial position errors and track biases, the most significant improvement in track forecasts is found at short lead-times within 96 hr, with the relative improvement decreasing as forecast lead-time increases. When the Monte Carlo resampling technique is applied to the ECMWF deterministic track forecasts from 2011, the calibrated track forecasts outperform the NHC forecasts on average by 10% beginning at a lead-time of 72 hr. One potential reason why this approach does not lead to greater improvement at longer lead-times is that the along-track and cross-track error

distribution for the 2010–2011 VarEPS forecasts is significantly different than the hindcasts. This result may indicate that at longer lead-times the hindcast track projections may not be representative of track biases in the real-time VarEPS. However, it may also suggest that a more sophisticated, dynamically based approach for track correction is needed. For instance, the track predictability of TCs that are generated in the eastern tropical Atlantic from easterly waves may exhibit unique forecast biases than TCs that are generated through the trough-induced pathway off the U.S. Atlantic coast. In addition, the track biases in the hindcasts and VarEPS may also be functionally dependent on the size, structure, and intensity of the model-derived vortex (Holland 1983). Hence, a dynamically based track calibration scheme that incorporates the differences in genesis location and pathway, and model-derived vortex structure and size may portend better track performance at extended forecast lead-times.

The statistical calibration approach used to improve intensity forecasts from the VarEPS is based on a quantile-to-quantile technique in which the distribution of bias-corrected ECMWF hindcast intensities is mapped to the distribution of observed TC intensities. Similar to the track calibration results, the intensity correction yields the largest improvement at shorter lead-times but also provides significant improvement relative to the uncorrected VarEPS forecasts at extended lead-times. However, this statistical technique cannot account for systematic model biases in the VarEPS due to incorrect parameterizations of physical processes or the lack of spatial resolution within the TC's inner core. Even with the new approach, calibrated intensity forecasts still do not reveal significant skill beyond 96 hr relative to climatology and persistence. A topic of future work is to utilize the logistic growth equation model, which is currently the best performing intensity forecast model in the North Atlantic, to generate an ensemble of LGEM forecasts from the VarEPS. Preliminary results by Musgrave et al. (2012) using a small ensemble of deterministic models indicate that an LGEM ensemble provides superior intensity guidance relative to a single deterministic LGEM forecast. The

working hypothesis is that a larger ensemble size that captures the variability in environmental characteristics for different ensemble forecast trajectories should provide a better state estimate for a TC's future intensity.

On intraseasonal time scales, the proposed forecast approach is to provide probabilistic guidance for the regional likelihood of TC activity during the next three to four weeks (Chapter 4 and Chapter 5). From this perspective, dynamically based forecasts from the ECMFS provide forecast skill exceeding climatology out to weeks three and four for the southern Gulf of Mexico, western Caribbean and the Main Development Region. Forecast skill in these regions is traced to the model's ability to capture correctly the variability in deep-layer vertical wind shear as well as the relative frequency of easterly waves moving through these regions.

In addition, the intraseasonal TC predictability study in Chapter 4 along with findings from its operational use in Chapter 5 indicates how one might predict the predictability of the ECMFS (i.e., Palmer 2000). The variance in predicted TC frequency, accuracy and reliability in the ECMFS is related strongly to the phase and amplitude of the MJO at the time of model initialization. For instance, when the ECMFS is initialized while the active convective phase of the MJO is located in the Indian Ocean, the ECMFS is likely to yield a basin-wide enhancement in TC activity relative to the model's climate state. The reliability and relative operating characteristic score will also be maximized when the model is initialized in this state suggesting higher than normal forecast confidence. The combination of bias-calibrated regional probabilistic forecast guidance along with objectively defined measures of forecast confidence should increase the utility of these intraseasonal TC forecasts for various end-users.

For longer time-scales, interannual tropical cyclone variability has been considered with respect to its relationship with easterly wave activity across the North Atlantic. A climatology of easterly waves has been developed for the East Pacific, tropical Atlantic, and Africa that utilizes several global reanalyses including CFS-R,

ERA-Interim, ERA-40, and NCEP/NCAR Reanalysis I. Using these reanalysis-derived easterly wave climatologies, the covariability between easterly wave frequency and intensification efficiency is examined relative to the annual number of TCs in the North Atlantic. A moderately positive and statistically significant relationship is seen, suggesting that approximately 20–30% of the variance in the number of TCs may be explained by the frequency of easterly waves. Although TC forecasts on interannual time scales are not developed in this dissertation, initial findings with regards to the easterly wave–tropical cyclone relationship in Chapter 6 suggest variability in easterly wave frequency is an additional predictability avenue to consider for forecasts on these time scales.

In addition, the relationship between easterly wave frequency and various large-scale climate modes has been considered, with the AMO and AMM exhibiting the strongest positive covariability with Atlantic easterly wave frequency. This relationship is manifested spatially with enhanced easterly wave genesis across the western Sahel, larger track density anomalies across the tropical Atlantic's MDR, and some suggestion of a northward shift in preferred zones of easterly wave tracks in the tropical Atlantic during positive AMO years. The magnitude of the relationship between easterly wave frequency and the NAO and PDO is weaker than with the AMO, and the relationship is inversely proportional, which implies that during positive NAO and PDO years, the number of easterly waves across the tropical Atlantic is on average lower than during negative NAO and PDO years. It is hypothesized that the NAO and PDO modulate easterly wave frequency in the tropical Atlantic through induced changes to the North Atlantic storm track, which has recently been shown to be as effective as localized convective heating over Africa at triggering easterly wave genesis (Leroux et al. 2011).

Besides changes in the number of easterly waves, the intensification efficiency of easterly waves is shown to have substantial interannual variability. In fact, the coefficient of variation, which is the ratio of the standard deviation to the mean, averages

50% for the four reanalyses over their respective period of record. In addition, the intensification efficiency of easterly waves has increased from 14% during the period 1960–1990 to 24% during the period 1995–2010. When variability in easterly wave intensification efficiency is considered relative to large-scale climate modes, a stronger relationship is present with the AMO and AMM than is seen with easterly wave frequency. In fact, the relationship is so robust that regressions of the AMO onto easterly wave frequency entirely reproduce the low-frequency trends that have been observed in both the ERA-Interim and CFS-R since 1980 and the NCEP/NCAR and ERA-40 since 1948 and 1958, respectively. This analysis offers a plausible physical explanation for the recent increase in the number of NATL TCs, as it has been concomitant with an increasing trend in both the number of tropical Atlantic easterly waves and easterly wave intensification efficiency. In addition, the easterly wave–tropical cyclone pathway is likely a more important mechanism governing how the AMO and AMM modulate North Atlantic TC frequency than previous thought (e.g., Thorncroft and Hodges 2001, Hopsch et al. 2007, Kossin and Vimont 2007). A topic of future work is to consider how large-scale environmental changes in sea surface temperature, column integrated diabatic heating, and environmental wind shear in association with these climate modes projects onto easterly waves in the tropical Atlantic.

The interannual variability in NATL TCs also leads to a number of sensible impacts in terms of TC damage, storm surge, coastal and inland flooding, and TC tornadoes. It is this last component, the variability in TC tornadoes that has been the subject of analysis in Chapter 7. To reconcile the inhomogeneous, historical TC tornado record, two statistical tornado models are developed using a set of a priori predictors for TC tornado formation. The models explain approximately 60–70% of the observed tornado frequency per Gulf TC landfall and 85–90% of the seasonal number of tornadoes, indicating they possess skill in forecasting the total magnitude of a tornado outbreak from a Gulf landfalling TC as well as their annual total. Reconstructions of annual tornado

activity based on these models show that the high number of TC tornadoes observed during 2004 and 2005 is unprecedented in the historical record since 1920. The synthetic TC tornado record also reflects decadal scale variations in association with the AMO. However, a comparison of the current warm phase of the AMO (1995–2008) with the previous warm phase period (1948–1964) shows that the median number of tornadoes per Gulf TC landfall has significantly increased, especially in association with landfalling TCs generating large tornado outbreaks. This change is likely a reflection of the increase in median TC size (by 35%) of Gulf landfalling TCs along with an increased frequency of large TCs at landfall.

A topic of future work is to utilize the TC tornado model along with calibrated TC tracks and other structural information from the real-time VarEPS to generate probabilistic TC tornado forecasts for U.S. landfalling TCs. The predictability of these forecasts could be independently assessed using the ECMWF hindcasts from 2010 and 2011 and may be compared with idealized reconstructions from the models in Chapter 7. In addition, the distribution of observed tornado formation from the TC center provides an opportunity to translate forecasts of TC tornado frequency into spatial probabilities of tornado exceedances, which would help quantify the local risk of tornadoes in association with a landfalling TC.

REFERENCES

Aberson, S., 1998: Five-day tropical cyclone track forecasts in the North Atlantic basin. *Wea. Forecasting*, **13**, 1005–1015.

Agudelo, P. A., C. D. Hoyos, J. A. Curry, and P. J. Webster, 2011: Probabilistic discrimination between large-scale environments of intensifying and decaying African easterly waves, *Clim. Dyn.*, **36**, 1379–1401.

Avila, L. A., and J. L. Guiney, 2000: Eastern North Pacific Hurricane Season of 1998. *Mon. Wea. Rev.*, **128**, 2990–3000.

Barkmeijer, J., R. Buizza, T. N. Palmer, K. Puri, and J.-F. Mahfouf, 2001: Tropical singular vectors computed with linearized diabatic physics. *Quart. J. Roy. Meteor. Soc.*, **127**, 685–708.

Barnes, L. R., D. M. Schultz, E. C. Gruntfest, M. H. Hayden, C. C. Benight, 2009: Corrigendum: false alarm rate or false alarm ratio? *Wea. Forecasting*, **24**, 1452–1454.

Barrett, B. S., and L. M. Leslie, 2009: Links between tropical cyclone activity and Madden-Julian Oscillation phase in the North Atlantic and Northeast Pacific basins. *Mon. Wea. Rev.*, **137**, 727–744.

Belanger, J. I., J. A. Curry, and C. D. Hoyos, 2009: Variability in tornado frequency associated with U.S. landfalling tropical cyclones, *Geophys. Res. Lett.*, **36**, L17805.

Belanger, J. I., J. A. Curry, and P. J. Webster, 2010: Predictability of North Atlantic tropical cyclones on intraseasonal time scales, *Mon. Wea. Rev.*, **138**, 4393–4401.

Belanger, J. I., P. J. Webster, J. A. Curry, and M. T. Jelinek, 2012: Extended prediction of North Indian Ocean tropical cyclones, *Wea. Forecasting*, In Press.

Berry, G. J., and C. Thorncroft, 2005: Case study of an intense African easterly wave. *Mon. Wea. Rev.*, **133**, 752–766.

Berry, G. J., C. Thorncroft, and T. Hewson, 2007: African easterly waves during 2004 – analysis using objective techniques. *Mon. Wea. Rev.*, **135**, 1251–1267.

Bessafi, M., and M. C. Wheeler, 2006: Modulation of South Indian Ocean tropical cyclones by the Madden-Julian Oscillation and convectively-coupled equatorial waves. *Mon. Wea. Rev.*, **134**, 638–656.

Beven, J. L. II, 1999: The boguscane – a serious problem with the NCEP medium range forecast model in the tropics. Preprints, *23rd Conf. on Hurricanes and Tropical Meteorology*, Dallas, TX, Amer. Meteor. Soc., 845–848.

Bister, M., and K. A. Emanuel, 1997: The genesis of Hurricane Guillermo: TEXMEX analyses and a modeling study. *Mon. Wea. Rev.*, **125**, 2662–2682.

Brier, G. W., 1950: Verification of forecasts expressed in terms of probability. *Mon. Wea. Rev.*, **78**, 1–3.

Brown, D. P., J. L. Franklin, and J. R. Rhome, 2008: Verification of the National Hurricane Center's experimental probabilistic tropical cyclone genesis forecasts. Preprints, *28th Conf. on Hurricanes and Tropical Meteorology*, Orlando, FL, Amer. Meteor. Soc.

Buizza, R., and T. N. Palmer, 1995: The singular vector structure of the atmospheric global circulation. *J. Atmos. Sci.*, **52**, 1434–1456.

Buizza, R., J. R. Bidlot, N. Wedi, M. Fuentes, M. Hamrud, G. Holt, and F. Vitart, 2007: The new ECMWF VAREPS (Variable Resolution Ensemble Prediction System). *Quart. J. Roy. Meteor. Soc.*, **133**, 681–695.

Buizza, R., M. Leutbecher, and L. Isaksen, 2008: Potential use of an ensemble of analyses in the ECMWF Ensemble Prediction System. *Quart. J. Roy. Meteor. Soc.*, **134**, 2051–2066.

Buizza, R., M. Leutbecher, L. Isaksen, and J. Haseler, 2010: The use of EDA perturbations in the EPS. *ECMWF Newsletter*, **123**, 22–27 [Available online at: http://ecmwf.int/publications/newsletters/pdf/123.pdf]

Burpee, R. W., 1972: The origin and structure of easterly waves in the lower troposphere of North Africa. *J. Atmos. Sci.*, **29**, 77–99.

Calheiros, R. V., and I. I. Zawadzki, 1987: Reflectivity-rain rate relationships for radar hydrology in Brazil. *J. Appl. Meteor.*, **26**, 118–132.

Camargo, S. J., M. C. Wheeler, and A. H. Sobel, 2009: Diagnosis of the MJO modulation of tropical cyclogenesis using an empirical index. *J. Atmos. Sci.*, **66**, 3061–3074.

Cangialosi, J. P., and J. L. Franklin, 2011: 2010 National Hurricane Center forecast verification report. NOAA/NWS/NCEP/National Hurricane Center, 77 pp. [Available online at www.nhc.noaa.gov/verification/pdfs/Verification_2010.pdf]

Cangialosi, J. P., and J. L. Franklin, 2012: 2011 National Hurricane Center forecast verification report. NOAA/NWS/NCEP/National Hurricane Center, 76 pp. [Available online at www.nhc.noaa.gov/verification/pdfs/Verification_2011.pdf]

Chan, J. C. L, and W. M. Gray, 1982: Tropical cyclone movement and surrounding flow relationships. *Mon. Wea. Rev.*, **110**, 1354–1374.

Charney, J. G., and A. Eliassen, 1964: On the growth of the hurricane depression. *J. Atmos. Sci.*, **21**, 68–75.

Charney, J. G., 1971: Geostrophic Turbulence, *J. Atmos. Sci.*, **28**, 1087–1094.

Chauvin, F., R. Roehrig, and J.-P. Lafore, 2010: Intraseasonal variability of the Saharan heat low and its link with mid-latitudes. *J. Climate*, **23**, 2544–2561.

Cho, J. Y. N., R. E. Newell, and J. D. Barrick, 1999: Horizontal wavenumber spectra of winds, temperature, and trace gases during the Pacific Exploratory Missions: 2. gravity waves, quasi-two-dimensional turbulence, and vortical modes. *J. Geophys. Res.*, **104**, 16,297–16,308.

Chu, J.-H., C. R. Sampson, A. S. Levin, and E. Fukada, 2002: The Joint Typhoon Warning Center tropical cyclone best tracks 1945–2000, Report, Joint Typhoon Warning Center, Pearl Harbor, Hawaii.

Craig, G. C., and S. L. Gray, 1996: CISK or WISHE as the mechanism for tropical cyclone intensification. *J. Atmos. Sci.*, **53**, 3528–3540.

Curtis, L., 2004: Midlevel dry intrusions as a factor in tornado outbreaks associated with landfalling tropical cyclones from the Atlantic and Gulf of Mexico, *Wea. Forecasting*, **19**, 411–427.

Davis, M. A. S., G. M. Brown, and P. Leftwich, 1984: A tropical cyclone data tape for the eastern and central North Pacific basins, 1949–1983: Contents, limitations, and uses, *NOAA Tech. Memo., NWS-NHC-25*, 16 pp.

Dee, D. P., and Coauthors, 2011, The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **137**, 553–597.

DeMaria, M., 1996: The effect of vertical shear on tropical cyclone intensity change. *J. Atmos. Sci.*, **53**, 2076–2087.

DeMaria, M., 1996: A history of hurricane forecasting for the Atlantic basin, 1920–1995. *Historical Essays on Meteorology 1919–1995*, J. R. Fleming, Ed., Amer. Meteor. Soc., 263–306.

DeMaria, M., J. A. Knaff, and B. H. Connell, 2001: A tropical cyclone genesis parameter for the tropical Atlantic. *Wea. Forecasting*, **16**, 219–233.

Dupont, T., M. Plu, P. Caroff, and G. Faure, 2011: Verification of ensemble-based uncertainty circles around tropical cyclone track forecasts, *Wea. Forecasting*, **26**, 664–676.

Elsberry, R. L., and L. E. Carr III, 2000: Consensus of dynamical tropical cyclone track forecasts—errors versus spread, *Mon. Wea. Rev.*, **128**, 4131–4138.

Elsberry, R. L., M. S. Jordan, and F. Vitart, 2011: Evaluation of the ECMWF 32-day ensemble predictions during 2009 season of western North Pacific tropical cyclone events on intraseasonal timescales. *Asia-Pacific J. Atmos. Sci.*, **47**, 305–318.

Emanuel, K. A., 1986: An air-sea interaction theory for tropical cyclones. Part 1: steady state maintenance. *J. Atmos. Sci.*, **43**, 585–604.

Emanuel, K. A., and D. S. Nolan, 2004: Tropical cyclone activity and global climate. Preprints, *26th Conf. on Hurricanes and Tropical Meteorology*, Miami, FL, Amer. Meteor. Soc., 240–241.

Epstein, E. S., 1969: A scoring system for probability forecasts of ranked categories. *J. Appl. Meteor.*, **8**, 985–987.

Ferranti, L., and S. Corti, 2011: New clustering products. *ECMWF Spring 2011 Newsletter*, **127**, 6–12.

Ferreira, R. N., W. H. Schubert, 1997: Barotropic aspects of ITCZ breakdown. *J. Atmos. Sci.*, **54**, 251–285.

Frank, N. L., 1975: Atlantic tropical systems of 1974. *Mon. Wea. Rev.*, **103**, 294–300.

Frank, W. M., 1977: The structure and energetics of the tropical cyclone I. storm structure, *Mon. Wea. Rev.*, **105**, 1119–1135.

Frank, W. M., and P. E. Roundy, 2006: The role of tropical waves in tropical cyclogenesis. *Mon. Wea. Rev.*, **134**, 2397–2417.

Frisch, U. 1995: *Turbulence*. Cambridge University Press, Cambridge.

Fontan, A.-C., and O. Cabanes, 2008: Applied research studies in progress at RSMC La Reunion -1- Microwave imagery database related to tropical cyclones -2- Estimation of tropical cyclone genesis probability over the South-West Indian Ocean with ECMWF ensemble prediction forecast. Preprints, *28th Conf. on Hurricanes and Tropical Meteorology*, Orlando, FL, Amer. Meteor. Soc.

Fu, X., and P.-C. Hsu, 2011: Extended-range ensemble forecasting of tropical cyclogenesis in the northern Indian Ocean: modulation of Madden-Julian Oscillation. *Geophys. Res. Lett.*, **38**, L15803, 1–4.

Gall, J. S., I. Ginis, S.-J. Lin, T. P. Marchok, J.-H. Chen, 2011: Experimental tropical cyclone prediction using the GFDL 25-km-resolution global atmospheric model. *Wea. Forecasting*, **26**, 1008–1019.

Gentry, R. C., 1983: Genesis of tornadoes associated with hurricanes, *Mon. Wea. Rev.,* **111**, 1793–1805.

Goerss, J., 2000: Tropical cyclone track forecasts using an ensemble of dynamical models. *Mon. Wea. Rev.*, **128**, 1187–1193.

Goldenberg, S. B., C. W. Landsea, A. M. Mestas-Nunez, and W. M. Gray, 2001: The recent increase in Atlantic hurricane activity: causes and implications. *Science*, **293**, 474–479.

Gray, W. M., 1968: Global view of the origin of tropical disturbances and storms. *Mon. Wea. Rev.*, **96**, 669–700.

Hagedorn, R., and Coauthors, 2012: Comparing TIGGE multi-model forecasts with reforecast-calibrated ECMWF ensemble forecasts. *ECMWF Research Department Technical Memorandum*, **663**, ECMWF, Shinfield Park, Reading RG2-9AX, UK, pp. 27 [Available online at http://www.ecmwf.int/publications/library/ecpublications/_pdf/tm/601-700/tm663.pdf]

Hamill, T. M., R. Hagedorn, and J. S. Whitaker, 2008: Probabilistic forecast calibration using ECMWF and GFS ensemble forecasts. Part II: precipitation. *Mon. Wea. Rev.*, **136**, 2620–2632.

Hart, R. E., 2003: A cyclone phase space derived from thermal wind and thermal asymmetry. *Mon. Wea. Rev.*, **131**, 585–616.

Hendricks, E. A., M. T. Montgomery, and C. A. Davis, 2004: On the role of "vortical" hot towers in formation of Tropical Cyclone Diana (1984). *J. Atmos. Sci.*, **61**, 1209–1232.

Hill, E. L., W. Malkin, and W. A. Schulz Jr., 1966: Tornadoes associated with cyclones of tropical origin – practical features, *J. Appl. Meteor.*, **5**, 745–763.

Hill, K. A., and G. M. Lackmann, 2009, Influence of environmental humidity on tropical cyclone size. *Mon. Wea. Rev.*, **137**, 3294–3315.

Hock, T. F., and J. L. Franklin, 1999: The NCAR GPS drop- windsonde. *Bull. Amer. Meteor. Soc.*, **80**, 407–420.

Holland, G. J., 1983: Tropical cyclone motion: environmental interaction plus a beta effect. *J. Atmos. Sci.*, **40**, 328–342.

Holland, G. J., J. Done, C. Bruyere, C. Cooper, and A. Suzuki-Parker, 2010: Model investigations of the effects of climate variability and change on future Gulf of Mexico tc activity. *Offshore Technology Conference*, **20690**, 1–13.

Holland, G. J., and J. Done, 2011: A global index for tropical cyclone damage potential. Preprints, *44th AGU Fall Meeting*, San Francisco, CA, Amer. Geophys. Union.

Hopsch, S. B., C. D. Thorncroft, K. Hodges, and A. Aiyyer, 2007: West African storm tracks and their relationship to Atlantic tropical cyclones. *J. Climate*, **20**, 2468–2483.

Hopson, T. M., and P. J. Webster, 2010: A 1–10-day ensemble forecasting scheme for the major river basins of Bangladesh: forecasting severe floods of 2003–07. *J. Hydromet.*, **11**, 618–641.

Hoyos, C. D. and P. J. Webster, 2007: The role of intraseasonal variability in the nature of Asian monsoon precipitation. *J. Climate*, **20**, 4402–4424.

HURDAT, 1920–2008, NOAA Hurricane Research Division of AOML, Miami, FL

Isaksen, L., and Coauthors, 2010: Ensemble of data assimilations at ECMWF. *ECMWF Research Department Technical Memorandum*, **636**, ECMWF, Shinfield Park, Reading RG2-9AX, UK, pp. 42 [Available online at http://www.ecmwf.int/publications/library/ecpublications/_pdf/tm/601-700/tm636.pdf]

Jarvinen, B. R., C. J. Neumann, and M. A. S. Davis, 1984: A tropical cyclone data tape for the North Atlantic basin, 1886–1983: Contents, limitations, and uses, *NOAA Tech. Memo., NWS-NHC-22*, 24 pp.

Johnson, C., and R. Swinbank, 2009: Medium-range multi-model ensemble combination and calibration. *Quart. J. Roy. Meteor. Soc.*, **135**, 777–794.

Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-year reanalysis project, *Bull. Amer. Meteor. Soc.*, **77**, 437–470.

Kaplan, J., M. DeMaria, and J. A. Knaff, 2010: A revised tropical cyclone rapid intensification index for the Atlantic and eastern North Pacific basins. *Wea. Forecasting*, **25**, 220–241.

Kimball, S. K., and M. S. Mulekar, 2004: A 15-year climatology of North Atlantic tropical cyclones. Part I: size parameters, *J. Climate*, **17**, 3555–3575.

Klotzbach, P. J., 2010: On the Madden-Julian Oscillation–Atlantic hurricane relationship. *J. Climate*, **23**, 282–293.

Kossin, J. P., and D. J. Vimont, 2007: A more general framework for understanding Atlantic hurricane variability and trends. *Bull. Amer. Meteor. Soc.*, **22**, 1767–1781.

Krishnamohan, K. S., K. Mohanakumar, and P. V. Joseph, 2012: The influence of Madden–Julian Oscillation in the genesis of North Indian Ocean tropical cyclones. *Theor. Appl. Climo.*, doi: 10.1007/s00704-011-0582-x.

Landsea, C., 2007: Counting Atlantic tropical cyclones back to 1900. *Eos, Trans. Amer. Geophys. Union*, **88**, 197–202.

Leroux, S., N. M. J. Hall, 2009: On the relationship between African easterly waves and the African easterly jet. *J. Atmos. Sci.*, **66**, 2303–2316.

Leroux, S., N. M. J. Hall, and G. N. Kiladis, 2011: Intermittent African easterly wave activity in a dry atmospheric model: influence of the extratopics. *J. Climate*, **24**, 5378–5396.

Leroy, A., and M. C. Wheeler, 2008: Statistical prediction of weekly tropical cyclone activity in the Southern Hemisphere, *Mon. Wea. Rev.*, **136**, 3637–3654.

Lilly, D. K., 1983: Stratified turbulence and the mesoscale variability of the atmosphere. *J. Atmos. Sci.*, **40**, 749–761.

Madden, R. A., and P. R. Julian, 1971: Detection of a 40–50 day oscillation in the zonal wind in the tropical Pacific. *J. Atmos. Sci.*, **28**, 702–708.

Majumdar, S. J., and P. M. Finocchio, 2010: On the ability of global ensemble prediction systems to predict tropical cyclone track probabilities. *Wea. Forecasting*, **25**, 659–680.

Maloney, E. D., and D. L. Hartmann, 2000a: Modulation of eastern North Pacific hurricanes by the Madden-Julian Oscillation. *J. Climate*, **13**, 1451–1460.

Maloney, E. D., and D. L. Hartmann, 2000b: Modulation of hurricane activity in the Gulf of Mexico by the Madden-Julian Oscillation. *Science*, **287**, 2002–2004.

Maloney, E. D., and J. Shaman, 2008: Intraseasonal variability of the West African monsoon and Atlantic ITCZ. *J. Climate*, **21**, 2898–2918.

Mann, M. E., K. Emanuel, G. J. Holland, and P. J. Webster, 2007: Atlantic tropical cyclones revisited. *EOS*, **88**, 349–350.

Mason, S. J., and N. E. Graham, 1999: Conditional probabilities, relative operating characteristics, and relative operating levels. *Wea. Forecasting*, **14**, 713–725.

Mason, S. J., 2004: On using "climatology" as a reference strategy in the Brier and ranked probability skill scores. *Mon. Wea. Rev.*, **132**, 1891–1895.

McBride, J. L., and R. Zehr, 1981: Observational analysis of tropical cyclone formation. Part II: comparison of non-developing versus developing systems. *J. Atmos. Sci.*, **38**, 1132–1151.

McCaul, E. W. Jr., 1991: Buoyancy and shear characteristics of hurricane-tornado environments. *Mon. Wea. Rev.*, **119**, 1954–1978.

McCaul, E. W. Jr., and M. L. Weisman, 1996: Simulations of shallow supercell storms in landfalling hurricane environments. *Mon. Wea. Rev.*, **124**, 408–429.

McTaggart-Cowan, R., G. D. Deane, L. F. Bosart, C. A. Davis, and T. J. Galarneau Jr., 2008: Climatology of tropical cyclogenesis in the North Atlantic (1948–2004). *Mon. Wea. Rev.*, **136**, 1284–1304.

McWilliams, J. C., 1984: The emergence of isolated coherent vortices in turbulent flow. *J. Fluid Mech.*, **146**, 21–43.

Miller, R. J., A. J. Schrader, C. R. Sampson, and T. L. Tsui, 1990: The automated tropical cyclone forecasting system (ATCF). *Wea. Forecasting*, **5**, 653–660.

Mirkin, B., 1996: *Mathematical Classification and Clustering*. Kluwer Academic, 428 pp.

Mo, K. C., 2000: The association between intraseasonal oscillations and tropical storms in the Atlantic basin. *Mon. Wea. Rev.*, **128**, 4097–4107.

Mohapatra, M., B. K. Bandyopadhyay, and A. Tyagi, 2011: Best track parameters of tropical cyclones over the North Indian Ocean: a review. *Natural Hazards*, doi: 10.1007/s11069-011-9935-0

Molinari, J., S. Skubis, D. Vollaro, F. Alsheimer, and H. E. Willoughby, 1998: Potential vorticity analysis of tropical cyclone intensification. *J. Atmos. Sci.*, **55**, 2632–2644.

Molinari, J., D. Vollaro, and K. L. Corbosiero, 2004: Tropical cyclone formation in a sheared environment: a case study. *J. Atmos. Sci.*, **61**, 2493–2509.

Montgomery, M. T., M. E. Nicholls, T. A. Cram, and A. B. Saunders, 2006: A vortical hot tower route to tropical cyclogenesis. *J. Atmos. Sci.*, **63**, 355–386.

Montgomery, M. T., N. V. Sang, R. K. Smith, and J. Persing, 2009: Do tropical cyclones intensify by WISHE? *Quart. J. Roy. Meteor. Soc.*, **135**, 1697–1714.

Musgrave, K. D., B. D. McNoldy, and M. DeMaria, 2012: Creation of a statistical ensemble for tropical cyclone intensity prediction, *30th Conf. on Hurricanes and Tropical Meteorology*, Ponte Vedra, FL, Amer. Meteor. Soc.

Neumann, C. J., B. R. Jarvinen, C. J. McAdie, and G. R. Hammer, 1999: *Tropical Cyclones of the North Atlantic Ocean*, 1871–1998, 206 pp., NOAA, Silver Spring, MD.

Ooyama, K., 1964: A dynamical model for the study of tropical cyclone development. *Geophys. Int.*, **4**, 187–198.

Palmer, T. N., and D. L. T. Anderson, 1994: The prospects for seasonal forecasting—A review paper. *Quart. J. Roy. Meteor. Soc.*, **120**, 755–793.

Palmer, T. N., 2000: Predicting uncertainty in forecasts of weather and climate. *Rep. Prog. Phys.*, **63**, 71–116.

Palmer, T. N., and Coauthors, 2009: Stochastic parameterization and model uncertainty. *ECMWF Research Department Technical Memorandum*, **598**, ECMWF, Shinfield Park, Reading RG2-9AX, UK, pp. 42 [Available online at: http://www.ecmwf.int/publications/library/ecpublications/_pdf/tm/501-600/tm598.pdf].

Pasch, R. J., and L. A. Avila, 1994: Atlantic tropical systems of 1992. *Mon. Wea. Rev.*, **122**, 539–548.

Pasch, R. J., P. A. Harr, L. A. Avila, J.-G. Jiing, and G. Elliot, 2006: An evaluation and comparison of predictions of tropical cyclogenesis by three global forecast models. Preprints, *27th Conference on Hurricanes and Tropical Meteorology*, Monterey, CA, Amer. Meteor. Soc.

Powell, M. D., and T. A. Reinhold, 2007: Tropical cyclone destructive potential by integrated kinetic energy. *Bull. Amer. Meteor. Soc.*, **88**, 513–526.

Puri, K., J. Barkmeijer, and T. N. Palmer, 2001: Ensemble prediction of tropical cyclones using targeted diabatic singular vectors. *Quart. J. Roy. Meteor. Soc.*, **127**, 709–734.

Rappaport, E. N., 2000: Loss of life in the United States associated with recent Atlantic tropical cyclones. *Bull. Amer. Meteor. Soc.*, **81**, 2065–2073.

Rappaport, E. N., and Coauthors, 2009: Advances and challenges at the National Hurricane Center. *Wea. Forecasting*, **24**, 395–419.

Rappaport, E. N., J.-G. Jiing, C. W. Landsea, S. T. Murillo, and J. L. Franklin, 2012: The Joint Hurricane Test Bed, it's first decade of tropical cyclone research-to-operations activities reviewed. *Bull. Amer. Meteor. Soc.*, **93**, 371–380.

Richardson, D., 2011: Changes to the operational forecasting system. *ECMWF Newsletter*, **130**, 1–40. [Available online at http://ecmwf.int/publications/newsletters/pdf/130.pdf]

Ritchie, E. A., and G. J. Holland, 1997: Scale interactions during the formation of Typhoon Irving. *Mon. Wea. Rev.*, **125**, 1377–1396.

Rivière, G. and I. Orlanski, 2007: Characteristics of the Atlantic storm-track eddy activity and its relation with the North Atlantic Oscillation. *J. Atmos. Sci.*, **64**, 241–266.

Saha, S., and Coauthors, 2010: The NCEP climate forecast system reanalysis. *Bull. Amer. Meteor. Soc.*, **91**, 1015–1057.

Saunders, A. and M. T. Montgomery, 2004: *A closer look at vortical hot towers within a tropical cyclogenesis environment*. Department of Atmospheric Science Blue Book, Colorado State University, Vol. 752, 105 pp.

Serra, Y. L., G. Kiladis, and M. Cronin, 2008: Horizontal and Vertical Structure of Easterly Waves in the Pacific ITCZ. *J. Atmos. Sci.*, **65**, 1266–1284.

Simmons, A. J., and Coauthors, 2007: Comparison of trends and low-frequency variability in CRU, ERA-40, and NCEP/NCAR analyses of surface air temperature. *J. Geophys. Res.*, **109**, D24115, 1–18.

Simmons, A. J., S. Uppala, D. Dee, and S. Kobayashi, 2007: The ERA interim reanalysis. ECMWF Newsletter, No. 110, ECMWF, Reading, United Kingdom, 25–35. [Available online at http://www.ecmwf.int/publications/newsletters/pdf/110_rev.pdf.]

Simmons, K. M, and D. Sutter, 2005: WSR-88D radar, tornado warnings, and tornado casualties, *Wea. Forecasting*, **20**, 301–310.

Simpson, J., E. A. Ritchie, G. J. Holland, J. Halverson, and S. Stewart, 1997: Mesoscale interactions in tropical cyclone genesis. *Mon. Wea. Rev.*, **125**, 2643–2661.

Simpson, R., 2012: Background and motivations for the Saffir / Simpson scale. Preprints, *30th Conf. on Hurricanes and Tropical Meteorology*, Ponte Vedra, FL, Amer. Meteor. Soc.

Smith, R. K., M. T. Montgomery, and N. V. Sang, 2009: Tropical cyclone spin-up revisited. *Quart. J. Roy. Meteor. Soc.*, **135**, 1321–1335.

Sriver, R., and M. Huber, 2006: Low frequency variability in globally integrated tropical cyclone power dissipation. *Geophys. Res. Lett.*, **33**, L11705, 1–5.

Thorncroft, C. D., and K. Hodges, 2001: African easterly wave variability and its relationship to Atlantic tropical cyclone activity. *J. Climate*, **14**, 1166–1179.

Toma, V. E., and P. J. Webster, 2010: Oscillations of the intertropical convergence zone and the genesis of easterly waves. Part I: diagnostics and theory., *Clim. Dyn.*, **34**, 587–604.

Toma, V. E., and P. J. Webster, 2010: Oscillations of the intertropical convergence zone and the genesis of easterly waves. Part II: numerical verification. *Clim. Dyn.*, **34**, 605–613.

Trenberth, K. E., and J. W. Hurrell, 1994: Decadal atmosphere-ocean variations in the Pacific. *Clim. Dyn*., **9**, 303–319.

Tsui, T., and R. Miller, 1988: Evaluation of western North Pacific tropical cyclone objective forecast aids. *Wea. Forecasting*, **3**, 76–85.

Uppala, S. M., and Coauthors, 2005: The ERA–40 reanalysis. *Quart. J. Roy. Met. Soc.*, **131**, 2961–3012.

U.S. Storm Events Database, 1950–2008, National Climatic Data Center, Asheville, NC

Verbout, S. M., D. M. Schultz, L. M. Leslie, H. E. Brooks, D. J. Karoly, and K. L. Elmore, 2007: Tornado outbreaks associated with landfalling hurricanes in the North Atlantic Basin: 1954–2004, *Meteor. Atmos. Phys.*, **97**, 255–271.

Vitart, F., J. L. Anderson, and W. F. Stern, 1997: Simulation of interannual variability of tropical storm frequency in an ensemble of GCM integrations. *J. Climate*, **10**, 745–760.

Vitart, F., and Coauthors, 2008: The new VAREPS-monthly forecasting system: A first step towards seamless prediction. *Quart. J. Roy. Meteor. Soc.*, **134**, 1789–1799.

Vitart, F., 2009: Impact of the Madden-Julian Oscillation on tropical storms and risk of landfall in the ECMWF forecast system. *Geophys. Res. Lett.*, **36**, L15802.

Webster, P. J. and C. Hoyos, 2004: Prediction of monsoon rainfall and river discharge on 15-30 day time scales. *Bull. Amer. Met. Soc.*, **85**, 1745–1765.

Webster, P. J., T. Hopson, C. Hoyos, A. Subbiah, H.-R. Chang, and R. Grossman, 2006: A three-tier overlapping predictability scheme: Tools for strategic and tactical decisions in the developing world. *Predictability of Weather and Climate*, T. N. Palmer and R. Hagedorn, Eds., Cambridge University Press, 645–673.

Webster, P. J., 2008: Myanmar's deadly daffodil. *Nature Geoscience*, **1**, 488–490.

Wheeler, M., and G. N. Kiladis, 1999: Convectively coupled equatorial waves: analysis of clouds and temperature in the wavenumber-frequency domain. *J. Atmos. Sci.*, **56**, 374–399.

Wheeler, M. C., and H. H. Hendon, 2004: An all-season real-time multivariate MJO index: Development of an index for monitoring and prediction. *Mon. Wea. Rev.*, **132**, 1917–1932.

WMO, 2011: Tropical cyclone operational plan for the Bay of Bengal and Arabian Sea, World Meteorological Organization Tech. Tropical Cyclone Programme Rep. TCP-21,WMO/TD-No. 84, 97 pp. [Available from World Meteorological Organization, Case Postale No. 5, HC-1211, Geneva 20, Switzerland; also online at http://www.wmo.int/pages/prog/www/tcp/documents/TCP-21_2011final.pdf.]

Wilks, D. S., 1995: *Statistical Methods in the Atmospheric Sciences*. Academic Press, 467 pp.

Wunsch, C., 2006: *Discrete Inverse and State Estimation Problems*, 377 pp., Cambridge University Press, New York, NY.

Zhu, H., W. Ulrich, and R. K. Smith, 2004: Ocean effects on tropical cyclone intensification and inner-core asymmetries. *J. Atmos. Sci.*, **61**, 1245–1258.